Global environmental predictors of benthic marine biogeographic structure

Christina L. Belanger[a,1,2], David Jablonski[a], Kaustuv Roy[b], Sarah K. Berke[a], Andrew Z. Krug[a], and James W. Valentine[c,2]

[a]Department of the Geophysical Sciences, University of Chicago, Chicago, IL 60637; [b]Section of Ecology, Behavior and Evolution, University of California at San Diego, La Jolla, CA 92093; and [c]Department of Integrative Biology and Museum of Paleontology, University of California, Berkeley, CA 94720

Contributed by James W. Valentine, July 20, 2012 (sent for review June 27, 2012)

Analyses of how environmental factors influence the biogeographic structure of biotas are essential for understanding the processes underlying global diversity patterns and for predicting large-scale biotic responses to global change. Here we show that the large-scale geographic structure of shallow-marine benthic faunas, defined by existing biogeographic schemes, can be predicted with 89–100% accuracy by a few readily available oceanographic variables; temperature alone can predict 53–99% of the present-day structure along coastlines. The same set of variables is also strongly correlated with spatial changes in species compositions of bivalves, a major component of the benthic marine biota, at the 1° grid-cell resolution. These analyses demonstrate the central role of coastal oceanography in structuring benthic marine biogeography and suggest that a few environmental variables may be sufficient to model the response of marine biogeographic structure to past and future changes in climate.

bivalves | climate | macroecology | sea-surface temperature

Biogeographic units (BUs), such as provinces or biomes, have been essential for understanding the macroecological and evolutionary processes underlying global biodiversity patterns (1–4) and are increasingly being incorporated in conservation planning (5–8). BUs are also becoming important from a global change perspective as components in models of biotic responses to global climate change in terrestrial settings (9). However, the factors that determine the biogeographic structure of benthic marine species in shallow-water habitats, where marine biodiversity is best documented and environmental changes are projected to be most severe (10–12), remain poorly understood, limiting our ability to model systems-level responses of marine biodiversity to environmental change. Here we use global datasets (Datasets S1, S2, S3, S4, and S5) to show that, in contrast to the more complex relationship between species richness and environment (13, 14), the biogeographic structure of coastal and continental shelf habitats can be predicted using just a few oceanographic parameters. This robust first-order link between specific environmental factors and large-scale biotic patterns establishes the importance of climate and oceanography in structuring marine faunas and provides important insights into modeling past and future responses of marine biogeography to global change.

We evaluate the correspondence between BUs and oceanographic variables using a model-fitting approach (*Materials and Methods*) to determine which oceanographic variables (individually or in combination) are most strongly correlated with biogeographic structure at both the ocean-basin and coastline scales (Table 1 and Tables S1 and S2). Our models focus on mean annual values and seasonal ranges of sea surface temperature, salinity, and productivity (hereafter TSP) because these variables have been previously hypothesized, separately or together, to affect taxonomic compositions and species richness in benthic marine systems (15–17). Biogeographic schemes for the world oceans differ mainly in the specific locations of several major provincial boundaries (1, 6, 8, 18–21). To reconcile such differences, we (*i*) use a recent synthesis of existing biogeographic schemes for coastal and shelf marine provinces (hereafter the Spalding scheme) as our general biogeographic framework (6) and (*ii*) test the robustness of our results using an independently derived biogeographic model (1) (hereafter the Valentine scheme) that was not used in the synthetic scheme (Fig. 1 *A* and *C*). We then use the taxonomic composition of bivalve mollusks, a heavily sampled and taxonomically standardized model system for marine biogeography (22, 23), to test whether (*i*) the correlation between spatial changes in oceanographic variables and bivalve species composition also holds at finer (1° grid cell) resolution and (*ii*) whether the spatial distribution of bivalve taxa can be used to reconstruct spatial patterns in oceanography.

Results and Discussion

Environmental Factors Predict BU Membership. For both of the biogeographic schemes used here, TSP predicts BU membership with 80–89% accuracy at the ocean-basin scale and >89% accuracy along individual coastlines (Table 1, Fig. 2, Figs. S1 and S2). The greater predictive power at the coastline level most likely reflects the a priori separation of spatially disjunct areas that share similar oceanographic conditions. Overall, the predictions of BUs along each coastline using TSP are qualitatively similar between biogeographic schemes, but the prediction accuracy is higher for the Valentine scheme on coastlines where it uses fewer BUs than does the Spalding scheme (Fig. 2, Table 1). Model fits for the multinomial logistic regressions on each coastline, assessed using pseudo-R^2 metrics (24–27), are strong (R^2 = 0.88–1.00) with the exception of the Indian Ocean coastline and the West Pacific Islands in the Valentine scheme (R^2 = 0.67 and 0.64, respectively), which are each spatially dominated by a single province. In all cases, multinomial logistic regressions using TSP correctly predicted province membership for more cells than would be expected by chance by a margin of 23–84% (likelihood-ratio χ^2 test, df = 6, $P < 0.001$; Tables S3 and S4). For all variables, the variance inflation factor is less than 5, indicating that multicollinearity among variables is unlikely to influence these analyses (Table S5) (28).

The few regions with low prediction accuracy warrant further analysis of biogeographic and oceanographic patterns (Fig. 1). Spalding BUs are least related to TSP in western India and northwestern Australia (Fig. 1*B*); Valentine BUs are least related to TSP in the Gulf of Mexico, eastern Australia, and southeastern United States (Fig. 1*D*). These mismatches are not randomly distributed; grid cells near province boundaries have

Author contributions: C.L.B., D.J., K.R., and J.W.V. designed research; C.L.B., D.J., S.K.B., and A.Z.K. performed research; C.L.B. and S.K.B. analyzed data; and C.L.B., D.J., K.R., S.K.B., A.Z.K., and J.W.V. wrote the paper.

The authors declare no conflict of interest.

[1]Present address: Department of Geology and Geological Engineering, South Dakota School of Mines and Technology, Rapid City, SD 57701.

[2]To whom correspondence may be addressed. E-mail: Christina.Belanger@sdsmt.edu or jwvsossi@socrates.berkeley.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1212381109/-/DCSupplemental.

Table 1. Multinomial logistic regression results using temperature, salinity, and productivity means and seasonal ranges (TSP) to predict biogeographic unit (BU) membership for two biogeographic schemes by coastline and by basin

	Spalding et al. (6)			Valentine (1)		
Region	AIC	% correct	No. of BUs	AIC	% correct	No. of BUs
Atlantic Basin	1574.2541	86.11	15	1554.5174	86.78	15
West Atlantic	369.4630	95.40	8	418.0608	94.92	9
East Atlantic	98.0001	100.00	8	150.0794	98.18	6
Pacific Basin	4950.5760	80.20	44	2874.1524	89.00	17
Indian Ocean	802.3963	89.66	17	124.2388	98.57	4
West Pacific	943.3720	91.48	19	206.1705	98.12	7
West Pacific Islands	543.0267	89.41	14	56.0002	100.00	5
East Pacific	129.4221	98.21	8	188.7168	97.14	9

AIC, Akaike information criterion. Percentage (%) correct is the number of 1° grid cells correctly classified into the correct BU. "No. of BUs" is the number of biogeographic units in each region. Salinity contributes more to prediction accuracy than productivity on all coastlines. Including salinity is more important for the Spalding scheme, especially in the Pacific and Indian Oceans.

higher proportions of misclassification in both of the biogeographic schemes used here (χ^2 test, $P < 0.001$, Table S6). Mismatches near BU boundaries are unlikely to simply result from poor sampling because at least some occur in well-sampled

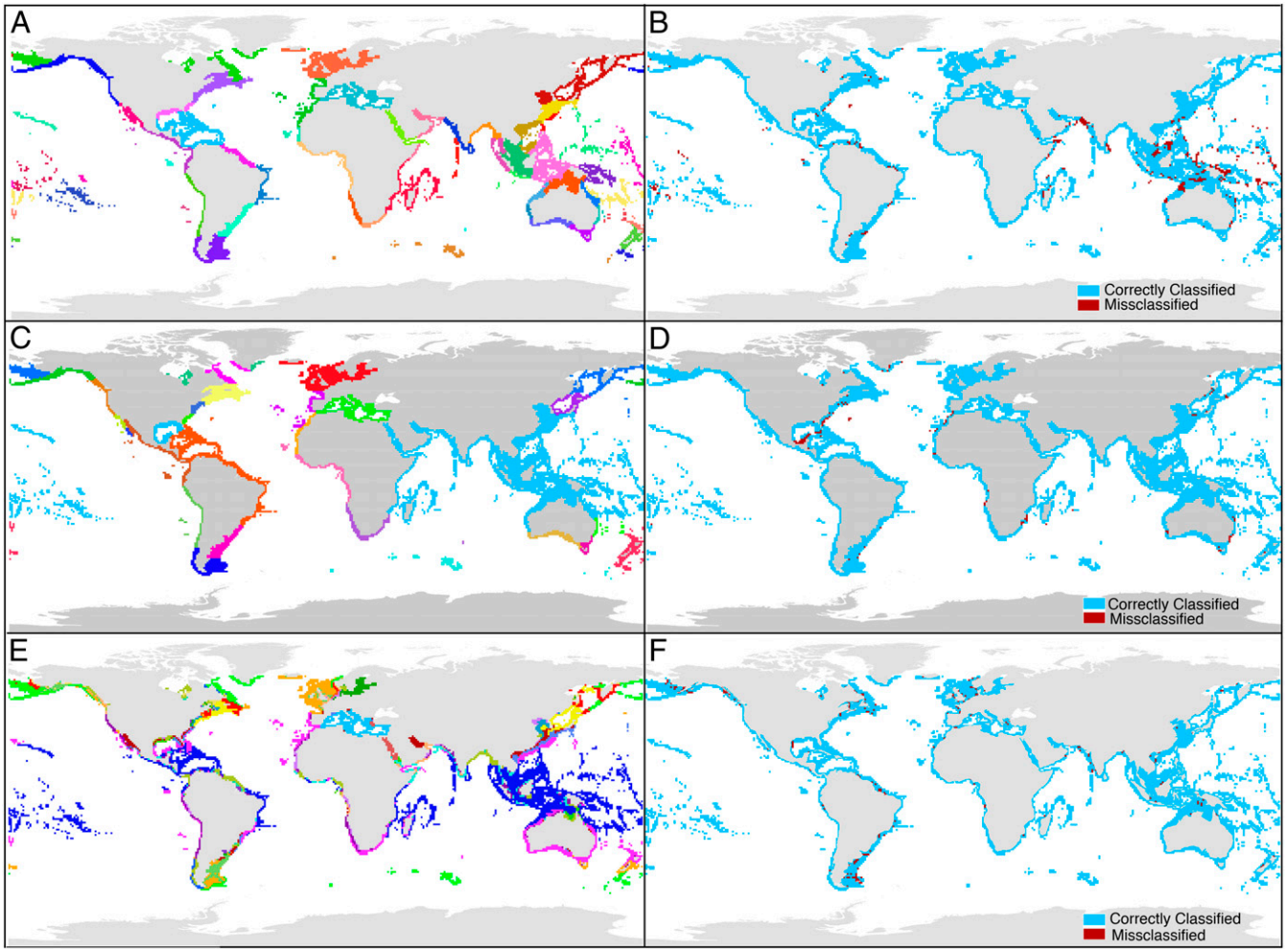

Fig. 1. Relationships between biogeographic structure and oceanographic conditions. (*A*) Spalding (6) biogeographic provinces, based principally on patterns of endemism. (*B*) Grid cells that are classified into the correct Spalding province using TSP (blue) and misclassified (red). (*C*) Valentine (1) biogeographic provinces, based principally on shared species and species diversity. (*D*) Grid cells that are classified into the correct Valentine province using TSP (blue) and misclassified (red). (*E*) Oceanographic units derived for a cluster analysis of TSP. (*F*) Grid cells that are classified into the correct oceanographic units using bivalve occurrence data (blue) and misclassified (red).

regions such as the southeastern United States and eastern Australia. They also do not consistently occur more frequently in grid cells with deeper benthic environments (χ^2 test, Table S6), suggesting that classification of cells is robust to the distance of the habitats from surface-water environmental predictors. In some cases, these mismatches correspond to boundaries that are included in the biogeographic scheme but are not recognized in the oceanographic structure (e.g., the Central Indo-Pacific; Fig. 1 *A*, *B*, and *E*). In other cases, it is the precise placement of the boundary that differs between BUs and oceanographic units, suggesting boundaries that undergo episodic changes in oceanographic conditions (e.g., the northwest Indian Ocean), which are not captured by the synoptic data used here. Temporal variability in oceanographic currents may increase the permeability of biogeographic boundaries, blurring the placement of BU boundaries due to taxonomic "leakage" (29). Alternatively, the misclassification could be affected by spatially variable environmental conditions near oceanographic boundaries that are not captured in the general placement of BU boundaries (Fig. 1*E*). For example, strong seasonality coupled with habitat heterogeneity along western Baja California del Sur results in a unique mixture of tropical and warm temperate molluscan faunas, but the region contains few endemics and thus is not a BU in the Spalding scheme. Despite these challenges, most BU boundaries are precisely predicted by TSP (Fig. 1).

Mean annual temperature emerges as the most important single environmental predictor of biogeographic units, although the full suite of oceanographic variables (TSP) is the best-supported model and has the highest accuracy on all coastlines except the Indian Ocean and West Pacific Islands in the Valentine scheme (Fig. 2, Table 2). Mean annual values of TSP are up to 30% more informative than seasonal ranges of these values alone (median = 17.5%). Along Pacific coastlines, adding salinity and productivity data improves BU prediction by 1–36% over annual means and seasonal ranges in temperature (T) (median improvement = 16.28%). In the Atlantic, however, TSP predicts BU membership only 4–10% more accurately than temperature alone (median improvement = 8.72%) (Fig. 2, Table 2). Therefore, although temperature is always the dominant variable in structuring shallow-marine biogeography, the relative importance of temperature varies among ocean basins and between province schemes. We ran models incorporating additional variables, including nutrient concentrations and oxygen saturation, but they do not consistently improve BU prediction over TSP and have less spatial coverage because of data limitations (*Materials and Methods*, Table S7). Factors such as biotic interactions or carbonate saturation states of seawater may also play a role in structuring benthic biogeography, but could not be tested here owing to the lack of synoptic data on a global scale.

Overall, these results suggest that the biogeographic structure of shallow water and continental shelf species is largely determined by the character of water masses and their boundaries, as has been reported for the surface waters of the open ocean (30–32). However, our analyses show that, in contrast to the pelagic realm (20), primary productivity plays a relatively minor role in structuring the biogeography of benthic organisms in the shallow oceans. The close coupling between oceanographic boundaries and biogeographic boundaries seen here could reflect physical barriers to dispersal owing to currents present at the boundary (29) as well as physiological challenges due to abrupt changes in environmental variables across such boundaries (33–35).

Environmental Factors Are Also Correlated with Bivalve Species Composition. If a limited set of oceanographic variables plays a major role in shaping the biogeography of the shallow-water biota, then finer-scale analyses should also detect correlations

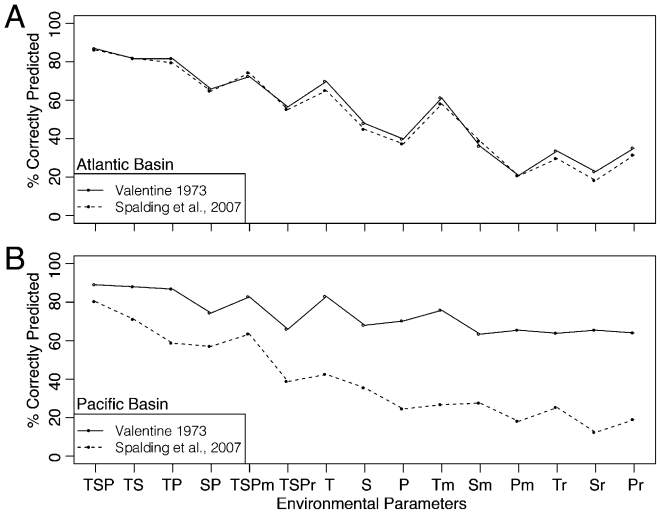

Fig. 2. Percentage of grid cells correctly predicted by nested multinomial logistic regressions where nested sets of oceanographic variables are used to predict BU membership at the basin scale for two biogeographic schemes. T, temperature; S, salinity; *P*, net primary productivity; m, mean annual values only; r, seasonal ranges only. Both mean annual values and seasonal ranges are used as predictors unless otherwise specified. See Tables S4 and S5 for corresponding AIC values and for coastline-scale analyses.

similar to those demonstrated here for relatively large BUs. We find such correlations at the 1° grid-cell resolution for coastal and shelf-depth bivalves. At this spatial scale, compositional changes in bivalve assemblages are strongly positively correlated with changes in TSP when the data are divided into Northern and Southern Hemisphere coastlines (Mantel test $r = 0.26$–0.77, $P < 0.001$; results are robust to choice of biotic distance metric; Table S8). These results indicate that the correlation between taxonomic composition and TSP holds both at the level of regional assemblages and at the level of large-scale biogeographic units.

Bivalve Biogeographic Structure Predicts Oceanographic Structure. Although oceanographic data can predict biogeographic structure, this does not demonstrate that biotic data can be used to reconstruct oceanographic structure, which would be valuable for analyses of coastal seas from the geological past where benthic fossil data are available but oceanographic data are scarce. To test whether bivalve taxonomic composition can predict oceanographic structure, we derived oceanographic units from a cluster analysis of TSP Euclidean distances among 1° grid cells; as temperature and salinity define water masses by determining water mass density, these factors reveal oceanographic structure (36, 37). We found that bivalve species compositions

Table 2. Improvement over temperature (T) data alone (annual means and seasonal range) in percent of one-degree grid-cells correctly classified when salinity (S) and productivity (P) data are added to analyses

	Spalding et al. (6)			Valentine (1)		
Region	TS	TP	TSP	TS	TP	TSP
West Atlantic	7.09	6.14	9.58	3.74	1.73	4.03
East Atlantic	6.49	3.42	7.86	8.09	6.84	9.91
Indian Ocean	25.42	14.31	30.59	3.63	3.15	3.63
West Pacific	21.60	6.17	24.48	5.58	3.14	5.98
West Pacific Islands	30.69	19.47	35.99	0.62	0.47	0.78
East Pacific	16.10	6.63	17.71	12.52	3.22	14.31

correctly predicted oceanographic unit membership with 98% accuracy in the total dataset at the maximum number of oceanographic clusters that was computationally feasible to analyze (*Materials and Methods*). Bivalve composition more accurately predicted oceanographic structure in the Pacific Basin than in the Atlantic Basin (Table 3). We also tested oceanographic configurations derived from the same cluster analysis that had fewer oceanographic units, and thus greater differences in TSP among the clusters, and found that bivalve biogeographic structure had higher prediction accuracy when there were fewer oceanographic units. The increase in accuracy with decreasing numbers of oceanographic units suggests that bivalve distributions are more likely to reflect large-scale oceanographic structure than smaller differences in oceanographic conditions. This likely arises in part from the range-through assumption used to determine bivalve occurrences at the 1° grid-cell resolution. This range-through protocol, necessary because the world's shelves are not evenly sampled at the 1° grid-cell scale, places bivalves in grid cells because their geographic range encompasses those grid cells, whether or not the species has been recorded from that location. However, even though the range-through assumption potentially overestimates occurrences of many species, bivalve compositions still reconstruct oceanographic units with a high degree of accuracy. This result suggests that the oceanographic structure of ancient shallow seas can be reconstructed from the biogeographic structure of fossil benthic macrofauna, a technique more commonly used with microfossil groups (31, 38, 39).

Taken together, the close correspondence between oceanographic and biogeographic structure documented here suggests that oceanographic conditions play a fundamental role in determining the large-scale biogeographic structure of the world oceans, with biotic interactions and other habitat characteristics playing a less direct role. Our results also suggest that biogeographic units do not represent obligate associations of taxa, but are composed of sets of species responding in similar ways to a limited set of environmental variables. Thus, as water masses and their boundaries are altered in response to global climate change [changes in temperature and salinity are already evident (40)], the biogeographic structure of benthic marine assemblages is also likely to change. The behavior of water masses under climate change—whether they expand, contract, split, or combine—will also have important evolutionary effects in terms of creating novel combinations of selective pressures within biogeographic units and erecting or eliminating barriers to gene flow between biogeographic units. Paleontological data from past warming events suggest that tropical water masses will expand (41–43), but the long-term consequence of this expansion to global diversity remains an open question. However, because the nature of the relationship between oceanographic and biogeographic structure varies among oceans, our results suggest that biotas from different ocean basins will respond differently to the same oceanographic forcing factors.

Table 3. Multinomial logistic regression results using bivalve species composition to predict oceanographic units (OUs)

Coastline/hemisphere	AIC	% correct	No. of OUs
West Atlantic/North	15259.1078	94.61	21
West Atlantic/South	4112.9569	85.43	10
East Atlantic/North	10463.5216	94.63	19
East Atlantic/South	776.3178	97.40	7
Indian Ocean	13617.1064	97.69	12
West Pacific/North	17425.0833	98.14	16
West Pacific/South	6732.5129	97.55	10
West Pacific Islands/North	768.0004	100.00	4
West Pacific Islands/South	1392.0002	100.00	4
East Pacific/North	7103.7724	91.36	16
East Pacific/South	1629.8629	96.21	9

AIC, Akaike information criterion. Percentage (%) correct is the number of 1° grid cells correctly classified into the correct OU. "No. of OUs" is the number of oceanographic units in each region.

Conclusions. Although correlations between temperature and large-scale spatial patterns in species richness of benthic marine taxa have been documented before (15, 44, but see ref. 13), few studies have directly quantified how oceanographic variables affect the geographic structure of the benthic marine biota. We demonstrate that a few readily acquired oceanographic parameters are sufficient to predict biogeographic patterns. Such information is critical not only for modeling biotic responses to future global climate change, but also for incorporating environmental factors into macroecological and evolutionary analyses of a major global system and for providing a more rigorous framework for interpreting paleobiogeographic data. The strong correlations between TSP and biogeographic structure documented here suggest that such systems-level changes may be predictable using model-based future projections of TSP. Our results also suggest that the distribution of well-preserved benthic fossil species should be useful for reconstructing the general paleoceanographic structure of ancient shallow seas, allowing a more rigorous application of data from past warm periods onto model predictions of shelf ecosystem change in a warm future.

Materials and Methods

Oceanographic Data and Data Coverage. The global ocean was divided into 1° grid cells, including only cells that contained (*i*) seafloor depths <200 m and (*ii*) mean annual and seasonal range data for sea-surface temperature, salinity, and net primary productivity. This protocol excluded areas with seasonal ice cover. Temperature and salinity data were obtained from the World Ocean Atlas 2009 quarter-degree grid-cell objectively averaged data set, which provides more complete coverage of coastal areas than the 1° data set (45–47). Quarter-degree cells were then averaged to achieve 1° resolution. Monthly averaged net primary productivity data were obtained from Oregon State University (48) at the one-sixth-degree cell resolution and averaged into 1° cells. Net primary productivity was calculated as a function of MODIS surface chlorophyll, MODIS sea-surface temperature, and MODIS cloud-corrected incident daily photosynthetically active radiation using the Vertically Generalized Production Model (48). This data set is available as a standard data product at http://www.science.oregonstate.edu/ocean.productivity/standard.product.php. For each variable, we calculated seasonal range as the difference between the maximum and minimum seasonal values. Annual means and seasonal ranges for TSP were available for 5,523 grid cells that could be assigned to biogeographic units and also contained data on the composition of the bivalve fauna (see below).

Assigning Grid Cells to Biogeographic Units. The province maps from Spalding et al. (6) were obtained from that paper's supplemental online materials (http://conserveonline.org) and province maps for Valentine (1) were manually digitized in ArcMap (49). Both were intersected with the 1° grid-cell template described above so that each grid cell had a provincial identity under each scheme. Grid cells that span neighboring provinces were assigned to the province representing the greatest area within that grid cell.

Predicting Biogeographic Structure from Environmental Variables. We used multinomial logistic regressions to test for correspondence between oceanographic conditions and province membership at the 1° grid-cell resolution using the function "multinom" in the "nnet" package in the R programming language (50, 51). In this procedure, we used oceanographic variables as predictors of province membership as determined by Spalding et al. (6) and Valentine (1). The percentage of grid cells classified into the correct biogeographic units measured the accuracy of predicting biogeographic units from oceanographic variables. Maximum-likelihood methods were used to determine the likelihood of those province predictions and are reported here as Akaike information criterion (AIC) values. We used these AIC values to compare nested subsets of oceanographic variables within province models to determine which variables are most associated with biogeographic structure. The proportion of cells expected to

be correctly classified by chance was calculated as the sum of the squared proportion of grid cells in each BU category in each region (52). Variance inflation factors were calculated using the "vif" function in the "HH" package in the R programming language (50, 53).

To further assess model fit, we used pseudo-R^2 metrics that compare a model using no predictors to a model using TSP as predictors. No one measure of R^2 is clearly a better measure of fit than another, so we calculated multiple estimators of pseudo-R^2 (24), using the "pR2" function in the "pscl" package of the R programming language (50, 54). Each method relies on taking two times the difference between the log-likelihood for the null (intercept only) multinomial logistic regression and the log-likelihood for the multinomial logistic regression using TSP as predictors (24).

Additional variables including dissolved oxygen, silicate, nitrate, and phosphate concentrations (55, 56) were available from the World Ocean Atlas 2009 for 3,565 1° grid cells used in this analysis, and these were also tested for their ability to predict BU membership with corresponding temperature, salinity, and productivity data as above.

Testing for Boundary Effects and Depth Effects in Mismatches Between TSP and BUs. To test whether a grid cell's proximity to a BU boundary contributed to its probability of being misclassified by the multinomial logistic regression, we coded grid cells by their distance from the nearest province boundary. Grid cells were then binned into two categories: "near" a province boundary and "away" from a province boundary for both biogeographic schemes. Distance cutoffs of 1°, 2°, and 3° were tested. We then tested whether the proportion of grid cells near a province boundary had a greater proportion of misclassifications than grid cells farther from the province boundary using χ^2 tests. We also tested whether shelf depth influenced grid-cell classification. Grid cells were binned into two categories, "shallow" and "deep," and three depth cutoffs were tested: 20, 50, and 100 m.

Correlations Between Bivalve Composition and Environmental Factors. We also tested whether the distribution of bivalve taxa, irrespective of clustering into provinces, is correlated with oceanographic patterns using Mantel tests between the dissimilarity in bivalve occurrences and the dissimilarity in oceanographic variables among 1° grid cells. Bivalves are a taxonomically and functionally diverse group with well-documented shallow-water distributions making them a model system for biogeography. We use a taxonomically standardized global database of occurrences for 5,460 species from 2,464 localities derived from the literature and from museum collections (22). Bivalve occurrences within each grid cell were determined by assuming that a given species is present in all grid cells contained within the total geographic range of the species. Thus, a species was counted as present wherever the box bounding its range intersected depths <200 m. Although this range-through assumption may cause species to be counted as present in cells where they do not actually occur, such errors would bias against finding strong correlations between bivalve distributions and oceanographic conditions (assuming that bivalve occurrences are in fact causally related to oceanographic conditions). Thus, this method is conservative for our purposes. We used both Jaccard and Simpson distance metrics to calculate dissimilarity in bivalve occurrences among grid cells. Oceanographic dissimilarity was calculated as the Euclidean distance between grid cells based on the *z*-score standardized means and seasonal ranges of temperature, salinity, and productivity.

For Mantel tests, we subdivided the data by coastline and then separated each coastline into northern and southern hemispheres (see Fig. S1 for coastline divisions). By analyzing such "hemisphere coastlines" we avoided comparing grid cells that have similar oceanographic conditions but are not connected by intervening seafloor at continental shelf depths and are thus difficult or impossible for species to disperse between. A strong positive correlation suggests that grid cells that were most similar in oceanographic variables were also most similar in their taxonomic composition. Both Pearson and Spearman correlations were calculated (Table S8).

Predicting Oceanographic Units from Bivalve Compositions. To define oceanographic units, we used a hierarchical cluster analysis based on the Euclidean distances of *z*-score standardized temperature, salinity, and productivity among 1° grid cells. Clusters were defined at successively increasing similarities in 0.5 intervals to create a nested hierarchy of potential oceanographic units. In each iteration, only existing clusters containing >5% of the total grid cells were allowed to be further split to prevent situations where clusters with large differences among constituent grid cells could be reduced to single-cell units (i.e., near river outflows) while geographically large areas with smaller differences remained intact (i.e., tropical seas). Cluster analyses were performed using the "hclust" function in the "stats" package of the R programming language (43). We then asked how well bivalve occurrence data could reconstruct the oceanographic units at each level of clustering using the multinomial logistic regression analyses described above until the number of clusters become too numerous to analyze (distance = 2.5, number of clusters = 24).

ACKNOWLEDGMENTS. We thank S. M. Kidwell and D. W. Bapst for valuable comments, S. Chiang for diligent assistance in data entry, and many bivalve systematists for advice and assistance, including L. C. Anderson, K. Amano, A. G. Beu, R. Bieler, J. G. Carter, R. von Cosel, J. S. Crampton, E. V. Coan, T. A., Darragh, H. H. Dijkstra, E. M. Harper, C. S. Hickman, S. Kiel, K. Lam, K. Lamprell, K. A. Lutaenko, N. Malchus, P. A. Maxwell, P. M. Mikkelsen, P. Middelfart, N. J. Morris, G. Paulay, A. Sartori, F. Scarabino, J. A. Schneider, P. V. Scott, J. T. Smith, J. D. Taylor, J. J. ter Poorten, J. D. Todd, T. R. Waller, A. Warén, and F. P. Wesselingh. This research was supported by National Science Foundation Grants EAR-0922156 and DEB-0919451 and National Aeronautics and Space Administration Grant EXOB08-0089.

1. Valentine JW (1973) *Evolutionary Paleoecology of the Marine Biosphere* (Prentice Hall, Englewood Cliffs, NJ), p 511.
2. Brown JH (1995) *Macroecology* (Univ of Chicago Press, Chicago), p 269.
3. Lomolino MV, Riddle BR, Brown JH (2006) *Biogeography* (Sinauer Associates, Sunderland, MA), p 845.
4. Witman JD, Roy K (2009) *Marine Macroecology* (Univ of Chicago Press, Chicago), p 424.
5. Whittaker RJ, et al. (2005) Conservation biogeography: Assessment and prospect. *Divers Distrib* 11:3–23.
6. Spalding MD, et al. (2007) A bioregionalization of coastal and shelf areas. *Bioscience* 57:573–583.
7. Abell R, et al. (2008) Freshwater ecoregions of the world: A new map of biogeographic units for freshwater biodiversity conservation. *Bioscience* 58:403–414.
8. Spalding MD, Agostini VN, Rice J, Grant SM (2012) Pelagic provinces of the world: a biogeographic classification of the world's surface pelagic waters. *Ocean Coast Manage* 60:19–30.
9. Sala OE, et al. (2000) Global biodiversity scenarios for the year 2100. *Science* 287: 1770–1774.
10. Gilbert D, Rabalais NN, Diaz RJ, Zhang J (2010) Evidence for greater oxygen decline rates in the coastal ocean than in the open ocean. *Biogeosciences* 7:2283–2296.
11. Belkin IM (2009) Rapid warming of large marine ecosystems. *Prog Oceanogr* 81: 207–213.
12. Lima FP, Wethey DS (2012) Three decades of high-resolution coastal sea surface temperatures reveal more than warming. *Nat Commun* 3:704.
13. Clarke A (2009) Temperature and marine macroecology. *Marine Macroecology*, eds Witman JD, Roy K (Univ of Chicago Press, Chicago), pp 250–278.
14. Mittelbach GG, et al. (2007) Evolution and the latitudinal diversity gradient: Speciation, extinction and biogeography. *Ecol Lett* 10:315–331.
15. Roy K, Jablonski D, Valentine JW, Rosenberg G (1998) Marine latitudinal diversity gradients: Tests of causal hypotheses. *Proc Natl Acad Sci USA* 95:3699–3702.
16. Gooday AJ, Jorissen FJ (2012) Benthic foraminiferal biogeography: Controls on global distribution patterns in deep-water settings. *Annu Rev Mar Sci* 4:237–262.
17. Somero GN (2012) The physiology of global change: Linking patterns to mechanisms. *Annu Rev Mar Sci* 4:39–61.
18. Forbes E (1856) Map of the distribution of marine life. *The Physical Atlas of Natural Phenonomena*, ed Johnston AK (William Blackwood and Sons, Edinburgh), pp 99–102; p. 131.
19. Briggs JC (1995) *Global Biogeography* (Elsevier, Amsterdam), p 452.
20. Longhurst A (1998) *Ecological Geography of the Sea* (Academic Press, San Diego), p 398.
21. Briggs JC, Bowen BW (2012) A realignment of marine biogeographic provinces with particular reference to fish distributions. *J Biogeogr* 29:12–30.
22. Rex MA, Crame JA, Stuart CT, Clarke A (2005) Large-scale biogeographic patterns in marine mollusks: A confluence of history and productivity? *Ecology* 86:2288–2297.
23. Krug AZ, Jablonski D, Valentine JW, Roy K (2009) Generation of Earth's first-order biodiversity pattern. *Astrobiology* 9(1):113–124.
24. Long JS (1997) *Regression Models for Categorical and Limited Dependent Variables* (Sage Publications, Thousand Oaks, CA), p 297.
25. McFadden D (1973) Conditional logit analysis of qualitative choice behavior. *Frontiers of Economics*, ed Zarembka P (Academic Press, New York), pp 105–142.
26. Maddala GS (1983) *Limited-Dependent and Qualitative Variables in Economics* (Cambridge Univ Press, Cambridge, UK).
27. Craig JG, Uhler R (1970) The demand for automobiles. *Can J Econ* 3:386–406.
28. Bowerman BL, O'Connell RT (1990) *Linear Statistical Models: An Applied Approach* (PWS-Kent Publishing Company, Boston), p 1024.
29. Gaylord B, Gaines SD (2000) Temperature or transport? Range limits in marine species mediated solely by flow. *Am Nat* 155:769–789.
30. Bradshaw JS (1959) Ecology of living planktonic Foraminifera in the North and Equatorial Pacific Ocean. *Cushman Foundation for Foraminiferal Research Contribution* 10:254–264.
31. Baumann K-H, Andruleit H, Bockel B, Geisen M, Kinkel H (2005) The significance of extant coccolithophores as indicators of ocean water masses, surface water temperature, and palaeoproductivity: A review. *Palaontologische Zeitschrift* 79:93–112.

32. Cermeño P, Falkowski PG (2009) Controls on diatom biogeography in the ocean. *Science* 325:1539–1541.
33. Valentine JW (1966) Numerical analysis of marine molluscan ranges on extratropical northeastern Pacific shelf. *Limnol Oceanogr* 11:198–211.
34. Hayden BP, Dolan R (1976) Coastal marine fauna and marine climates of the Americas. *J Biogeogr* 3:71–81.
35. Roy K, Jablonski D, Valentine JW (1995) Thermally anomalous assemblages revisited: Patterns in the extraprovincial latitudinal range shifts in Pleistocene marine mollusks. *Geology* 23:1071–1074.
36. Sverdrup HU, Johnson MW, Fleming RH (1942) *The Oceans: Their Physics, Chemistry, and General Biology* (Prentice-Hall, New York), p 1087.
37. Emery WJ (2003) Water types and water masses. *Encyclopedia of Atmospheric Sciences*, eds Holton JR, Curry JA, Pyle JA (Elsevier), 2nd Ed, pp 1556–1567.
38. Denne RA, Sen Gupta BK (1993) Matching of benthic foraminiferal depth limits and water-mass boundaries in the northwestern Gulf of Mexico: An investigation of species occurrences. *J Foraminiferal Res* 23:108–117.
39. Murray JW (1995) Microfossil indicators of ocean water masses, circulation, and climate. *Geol Soc Lond Spec Publ* 83:245–264.
40. Bindoff NL, et al. (2007) Observations: oceanic climate change and sea level. *Contribution of Working Group I to the Forth Assessment Report of the Intergovernmental Panel on Climate Change*, eds Solomon S, et al. (Cambridge Unive Press, Cambridge, UK), pp 387–432.
41. Oleinik A, Marincovich L, Jr., Barinov KB, Swart PK (2008) Magnitude of Middle Miocene warming in North Pacific high latitudes: Stable isotope evidence from Kaneharaia (Bivalvia, Dosiniinae). *Bulletin of the Geological Survey of Japan* 59:339–353.
42. Dowsett HJ, Chandler MA, Robinson MM (2009) Surface temperatures of the Mid-Pliocene North Atlantic Ocean: Implications for future climate. *Philos Transact A Math Phys Eng Sci* 367(1886):69–84.
43. Lutz BP (2011) Shifts in North Atlantic planktic foraminifer biogeography and subtropical gyre circulation during the mid-Piacezian warm period. *Mar Micropaleontol* 80:125–149.
44. Tittensor DP, et al. (2010) Global patterns and predictors of marine biodiversity across taxa. *Nature* 466:1098–1101.
45. Boyer T, et al. (2005) Objective analyses of annual, seasonal, and monthly temperature and salinity for the world ocean on a 0.25O grid. *Int J Climatol* 25: 931–945.
46. Antonov JI, et al. (2010) World Ocean Atlas 2009, Vol. 2: Salinity. *NOAA Atlas NESDIS 69*, ed Levitus S (US Government Printing Office, Washington DC), p 184.
47. Locarnini RA, et al. (2010) World Ocean Atlas 2009, Vol. 1: Temperature. *NOAA Atlas NESDIS 68*, ed Levitus S (US Government Printing Office, Washington, DC), p 184.
48. Behrenfeld MJ, Falkowski PG (1997) Photosynthetic rates derived from satellite-based chlorophyll concentration. *Limnol Oceanogr* 42:1–20.
49. ESRI (2011) *ArcGIS Desktop: Release 10* (Environmental Systems Research Institute, Redlands, CA).
50. R Development Core Team (2010) *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, Vienna).
51. Venables WN, Ripley BD (2002) *Modern Applied Statistics with S* (Springer, New York), 4th Ed, p 512.
52. Petrucci CJ (2009) A primer for social worker researchers on how to conduct multinomial logistic regression. *J Soc Serv Res* 35:193–205.
53. Heiberger RM (2009) *HH: Statistical Analysis and Data Display: Heilberger and Holland*. R package version 2.1-32.
54. Jackman S (2010) *pscl: Classes and Methods for R Developed in the Political Science Computational Laboratory* (Stanford). R package version 1.06.6.
55. Garcia HE, et al. (2010) World Ocean Atlas 2009, Vol. 4: Nutrients (phosphate, nitrate, silicate). *NOAA Atlas NESDIS 71*, ed Levitus S (US Government Printing Office, Washington, DC), p 398.
56. Garcia HE, et al. (2010) World Ocean Atlas 2009, Vol. 3: Dissolved Oxygen, Apparent Oxygen Utilization, and Oxygen Saturation. *NOAA Atlas NESDIS 70*, ed Levitus S (US Government Printing Office, Washington, DC), p 344.

Supporting Information

Belanger et al. 10.1073/pnas.1212381109

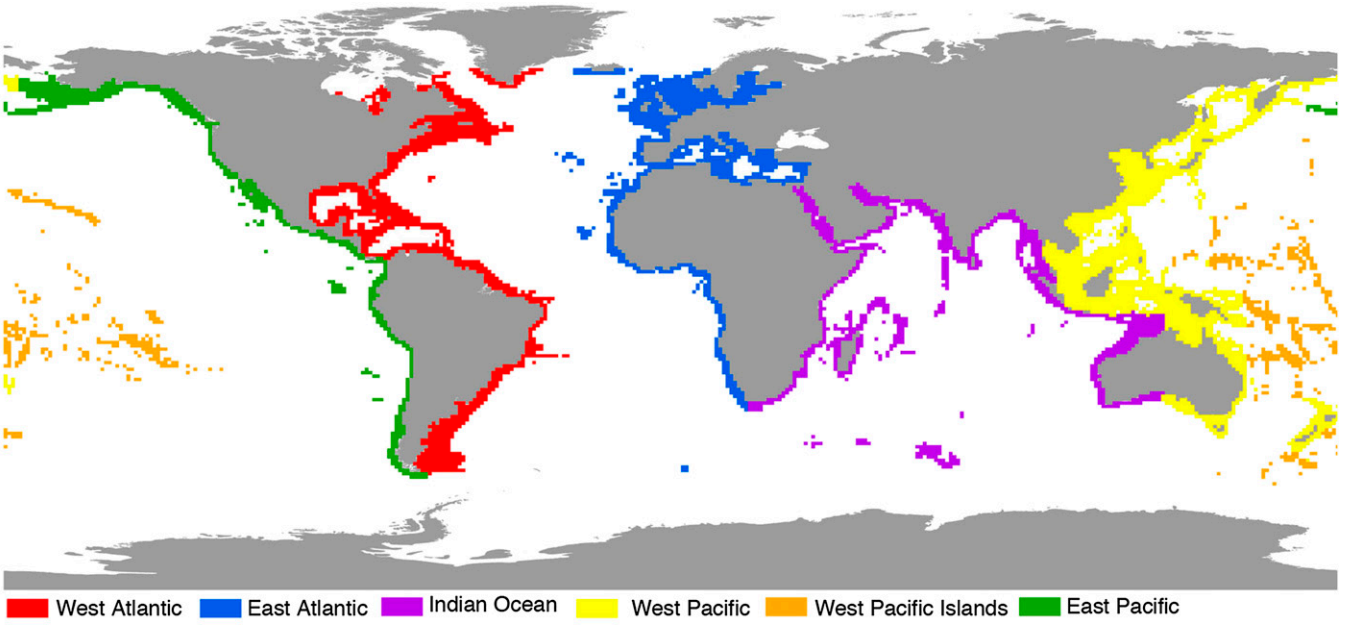

Fig. S1. Map of coastline assignments used for analysis.

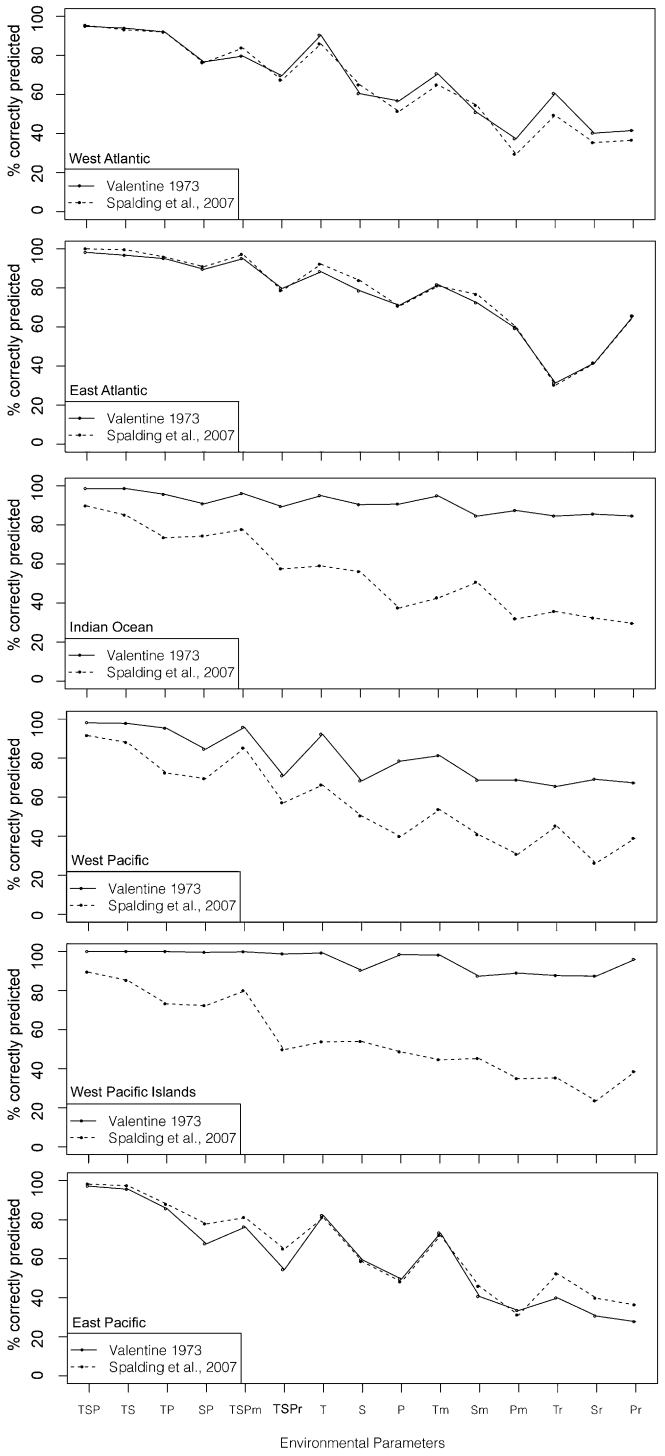

Fig. S2. Proportion of grid cells correctly predicted with nested multinomial logistic regressions where nested sets of oceanographic variables are used to predict biogeographic units membership at the coastline-scale for two biogeographic schemes. T, temperature; S, salinity; *P*, net primary productivity; m, mean annual values only; r, seasonal ranges only. Both mean annual values and seasonal ranges are used as predictors unless otherwise specified.

Table S1. Results from multinomial logistic regressions at the basin scale for combinations of temperature, salinity, and productivity

		Spalding et al. (6)		Valentine (1)	
Region	Oceanographic variables	AIC	% correct	AIC	% correct
Atlantic Basin	TSP	1574.2541	86.11	1554.5174	86.78
	TS	1919.0188	81.79	2019.5277	81.63
	TP	2310.6245	79.45	2077.5940	81.63
	SP	4028.2070	64.72	4119.2604	65.87
	Mean	2889.1195	74.19	2872.6364	72.32
	Range	5018.7010	55.20	5084.2548	56.56
	T	3253.0357	64.98	2995.8055	69.77
	S	5462.3780	44.80	5998.9693	47.92
	P	6853.6874	37.36	6642.2616	39.85
	Tm	4450.0939	58.06	4085.2605	61.19
	Sm	6520.2745	38.76	7195.8872	36.06
	Pm	8553.6308	20.71	8453.4566	20.92
	Tr	7274.3786	29.66	7183.0434	33.61
	Sr	8494.1120	18.31	8668.4117	22.89
	Pr	7897.0784	31.48	7886.0499	35.12
Pacific Basin	TSP	4950.5760	80.20	2874.1524	89.00
	TS	6827.1300	70.98	3233.8184	87.95
	TP	9671.5030	58.65	3453.5118	86.75
	SP	10274.8327	56.98	6522.6356	74.31
	Mean	8423.5462	63.40	4312.6558	82.70
	Range	14514.1439	38.74	8245.1403	65.90
	T	13370.2547	42.43	4096.7124	82.92
	S	15326.5947	35.46	8133.6998	67.95
	P	18190.9774	24.52	8470.1694	70.17
	Tm	16832.9861	26.74	5643.4275	75.76
	Sm	17668.5572	27.55	9607.9787	63.34
	Pm	21137.3089	18.11	10377.8202	65.40
	Tr	19529.4645	25.27	9779.9474	63.82
	Sr	22653.6593	12.36	9891.5673	65.40
	Pr	20505.8611	18.94	10009.5586	63.98

Unmodified abbreviations signify combined annual mean and seasonal range; m and r signify annual mean or seasonal range alone, respectively. TSP, temperature, salinity, and productivity; AIC, Akaike information criterion.

Table S2. Results from multinomial logistic regressions at the coastline scale for combinations of temperature, salinity, and productivity

	East Pacific		West Pacific Islands		West Pacific		Indian Ocean		East Atlantic		West Atlantic	
Oceanographic variables	% correct	AIC	% correct	AIC	% correct	AIC	% correct	AIC	% correct	AIC	% correct	AIC
Spalding biogeographic scheme												
TSP	98.21	129.4221	89.41	543.0267	91.48	943.3720	89.66	802.3963	100.00	98.0001	95.40	369.4630
TS	97.32	136.0872	85.20	640.4002	87.99	1167.9600	84.93	1072.2375	99.54	96.5202	93.10	462.6210
TP	87.84	381.9150	73.21	1086.9563	72.37	2278.2423	73.38	1504.0063	95.79	278.4827	91.95	545.7276
SP	77.82	677.8696	72.27	1028.6839	69.48	2744.6094	74.26	1614.6819	90.89	624.7677	76.25	1648.8195
Mean	81.04	545.1232	79.91	746.0326	85.18	1304.6531	77.56	1287.4154	97.15	229.9975	83.81	1052.5928
Range	64.94	1118.6570	49.69	1725.3948	57.14	3606.1079	57.43	2328.2095	78.59	1241.1131	67.34	1888.1514
T	81.04	514.3560	53.74	1545.6950	66.20	2857.3405	58.97	2208.7251	92.14	379.2801	85.92	721.0794
S	58.50	1056.3590	53.89	1594.5792	50.37	4016.4624	56.00	2264.3980	83.71	927.0010	64.94	2288.6199
P	48.12	1369.7961	48.60	1800.8348	39.84	5434.0121	37.40	3502.4659	70.62	1649.2970	51.34	2725.8336
Tm	71.91	747.9727	44.55	1717.8045	53.66	3731.3913	42.57	3067.9530	80.98	815.0344	64.94	1576.4913
Sm	45.80	1542.1934	45.17	1858.2543	40.78	4727.7421	50.61	2726.5602	76.65	1246.1797	54.41	2915.8247
Pm	31.13	1836.2629	34.89	2283.5388	30.72	6598.5547	31.79	4017.9138	59.91	2238.9964	29.41	3742.7940
Tr	52.24	1464.1014	35.20	2312.7968	45.20	4758.8200	35.64	3567.7446	30.18	2223.1635	49.33	2571.8793
Sr	39.71	1804.2198	23.52	2760.9495	26.16	6546.0420	32.23	3869.2800	41.34	2447.5293	35.34	3505.6058
Pr	36.31	1788.3689	38.47	2222.9836	38.90	5986.0386	29.48	4064.4912	65.38	1987.9158	36.59	3399.5545
Valentine biogeographic scheme												
TSP	99.46	123.6955	100.00	56.0002	98.57	124.2388	98.12	206.1705	98.18	150.0794	94.92	418.0608
TS	97.67	156.0741	100.00	41.6876	98.57	122.2391	97.79	216.9852	96.70	193.7719	93.97	475.8171
TP	88.37	358.0478	100.00	42.4404	95.60	204.2373	95.31	427.4692	94.99	258.6252	92.05	533.4627
SP	71.56	820.9125	99.53	66.5420	90.76	479.0003	84.57	1307.2501	89.52	556.6310	76.72	1649.7534
Mean	82.11	507.9120	99.84	38.8196	96.04	219.8138	95.77	326.9835	94.99	276.5889	79.69	1101.2957
Range	55.81	1545.4198	98.75	84.4317	89.33	567.0904	70.96	2387.8227	79.73	1195.0348	69.64	1935.9592
T	85.15	412.0605	99.22	53.4258	94.94	241.0578	92.22	647.7484	88.27	486.5286	90.42	630.8623
S	60.47	1152.8282	90.34	427.3278	90.32	556.4764	68.34	2232.0852	78.47	992.0194	60.44	2687.8897
P	51.70	1506.1099	98.44	96.7516	90.65	654.2027	78.47	2353.1800	71.07	1606.9679	56.70	2514.8156
Tm	76.74	613.0932	98.13	84.3949	94.83	236.5041	81.29	1398.9948	81.66	868.9825	70.59	1341.4671
Sm	41.32	1704.7456	87.38	559.5594	84.49	1075.5462	68.61	2378.7077	72.32	1322.2342	50.77	3395.5260
Pm	33.81	1951.9892	88.94	448.6964	87.35	841.4309	68.75	3307.4376	59.11	2175.1007	37.36	3654.5939
Tr	43.29	1869.0942	87.69	523.6450	84.49	1050.2493	65.53	2865.7792	31.32	2210.9549	60.54	2455.8221
Sr	31.84	1915.1965	87.38	565.5770	85.48	761.2677	69.15	2986.9512	41.57	2462.5062	40.23	3676.5885
Pr	30.23	2005.7643	95.95	149.7016	84.49	1053.8006	67.27	3130.0102	65.72	1964.9854	41.57	3358.8725

Unmodified abbreviations signify combined annual mean and seasonal range; m and r signify annual mean or seasonal range alone, respectively. AIC, Akaike information criterion; TSP, temperature, salinity, and productivity.

Table S3. Percentage improvement in biogeographic units prediction accuracy by temperature, salinity, and productivity over the expectation by chance

	% correct by chance		% correct using TSP		Improvement over chance	
Region	Spalding	Valentine	Spalding	Valentine	Spalding	Valentine
Atlantic Basin	10.49	11.80	86.11	86.78	75.62	74.98
East Atlantic	27.16	27.08	95.40	94.92	68.24	67.84
West Atlantic	16.29	20.96	100.00	98.18	83.71	77.22
Pacific Basin	4.05	44.04	80.20	89.00	76.15	44.96
Indian Ocean	9.46	72.22	89.66	98.57	80.20	26.35
West Pacific	12.25	49.72	91.48	98.12	79.23	48.40
West Pacific Islands	12.50	77.10	89.41	100.00	76.91	22.90
East Pacific	19.85	17.12	98.21	97.14	78.36	80.02

Table S4. Goodness-of-fit of the multinomial logistic regressions and the effect of temperature, salinity, and productivity on predicting biogeographic unit (BU) membership

	Likelihood-ratio statistic		McFadden's R^2		Maximum likelihood R^2		Maximum likelihood Craig and Uhler R^2	
Region	Spalding	Valentine	Spalding	Valentine	Spalding	Valentine	Spalding	Valentine
East Atlantic	3780.44	3544.67	0.93	0.92	0.97	0.97	0.99	0.99
West Atlantic	2651.00	2562.21	1.00	0.97	0.95	0.95	1.00	1.00
Indian Ocean	4100.06	1005.55	0.88	0.92	0.99	0.67	0.99	0.96
West Pacific	6473.60	3170.69	0.90	0.96	0.99	0.88	1.00	0.99
West Pacific Islands	2457.21	650.19	0.88	0.99	0.98	0.64	0.99	0.99
East Pacific	1976.99	2031.10	0.98	0.96	0.97	0.97	1.00	1.00

All likelihood-ratio statistics are significant ($P << 0.001$; χ^2 test, df = 6).

Table S5. Variance inflation factors (VIF) for the *z*-score–transformed oceanographic variables (temperature, salinity, and productivity) used as predictors in multinomial logistic regressions

Variable	Mean annual temperature	Seasonal range temperature	Mean annual salinity	Seasonal range salinity	Mean annual productivity	Seasonal range productivity
VIF	1.62	1.64	2.21	1.29	3.13	4.76

Table S6. χ^2 tests comparing the proportion of misclassified grid cells with respect to proximity to biogeographic unit (BU) boundaries and sea floor depth for the two biogeographic schemes

	Classified correctly	Misclassified	% correct	χ^2 statistic	*P* value
Spalding biogeographic scheme					
Proximity to BU boundary					
Within 1°	542	132	80.42		
Outside 1°	4,624	225	95.36	216.1219	$P << 0.001$
Within 2°	1,148	200	85.16		
Outside 2°	4,018	157	96.24	204.943	$P << 0.001$
Within 3°	1,732	251	87.34		
Outside 3°	3,434	106	97.01	194.7066	$P << 0.001$
Depth of sea floor					
Shallower than or equal to 20 m	1,932	100	95.08		
Deeper than 20 m	3,234	257	92.64	12.4347	$P < 0.001$
Shallower than or equal to 50 m	2,479	126	95.16		
Deeper than 50 m	2,687	231	92.08	21.0817	$P < 0.001$
Shallower than or equal to 100 m	2,886	150	95.06		
Deeper than 100 m	2,280	207	91.68	25.3152	$P < 0.001$
Valentine biogeographic scheme					
Proximity to BU boundary					
Within 1°	245	53	82.21		
Outside 1°	5,146	79	98.49	313.0876	$P << 0.001$
Within 2°	565	83	87.19		
Outside 2°	4,826	49	98.99	336.5482	$P << 0.001$
Within 3°	888	103	89.61		
Outside 3°	4,503	29	99.36	327.4437	$P << 0.001$
Depth of sea floor					
Shallower than or equal to 20 m	1,979	53	97.39		
Deeper than 20 m	3,412	79	97.74	0.5168	$P = 0.4722$
Shallower than or equal to 50 m	2,543	62	97.62		
Deeper than 50 m	2,848	70	97.60	0.0018	$P = 0.9662$
Shallower than or equal to 100 m	2,967	69	97.73		
Deeper than 100 m	2,424	63	97.47	0.2937	$P = 0.5879$

Three distance thresholds (1°, 2°, or 3°) are tested for the effect of grid-cell proximity to BU boundaries. Significant values indicate proportionally more misclassification in grid cells near BU boundaries. Three depth thresholds (20, 50, and 100 m) were tested for the effect of sea floor depth. Significant values indicate proportionally more misclassification in grid cells near BU boundaries. In the Valentine scheme, there is no significant difference in misclassification with depth.

Table S7. Results from multinomial logistic regressions at the coastline level for combinations of environmental variables

	West Atlantic		East Atlantic		Indian Ocean		West Pacific		West Pacific Islands		East Pacific	
Region	AIC	% correct	AIC	% correct	AIC	% correct	AIC	% correct	AIC	% correct	AIC	% correct
Spalding biogeographic scheme												
TSONP	210.0002	100.00	210.0002	100.00	480.0002	100.00	540.0034	100.00	360.0001	100.00	210.0001	100.00
TSON	182.0001	100.00	182.0002	100.00	416.0001	100.00	531.0357	98.80	312.0001	100.00	182.0001	100.00
TSO	188.2927	97.54	98.0002	100.00	549.6126	89.58	700.7716	89.08	492.8153	89.34	121.0844	97.79
TSN	154.0137	100.00	154.0001	100.00	425.1023	97.44	608.5806	96.40	356.7030	97.63	154.0002	100.00
TSP	173.5382	97.54	98.0001	100.00	429.0200	94.15	634.8296	92.47	502.8801	89.17	113.3298	98.74
TON	172.9079	99.57	154.0002	100.00	519.6652	94.15	922.8670	88.32	465.2104	96.45	154.0001	100.00
TOP	255.4465	96.09	99.6271	100.00	547.7375	89.76	1234.5580	74.56	929.7474	75.30	159.3552	96.21
TNP	173.1042	99.13	154.0001	100.00	382.8483	98.90	901.0330	89.08	285.3134	99.32	154.0001	100.00
TS	208.5642	95.95	70.0004	100.00	632.3840	86.84	791.8093	86.24	584.6978	84.77	111.1418	96.85
TO	352.0529	92.04	179.8994	95.95	919.5690	74.59	1586.8344	67.69	1274.0108	62.44	213.7364	91.48
TN	171.0120	98.70	126.0001	100.00	609.4069	88.85	1126.6378	84.17	529.6888	92.22	126.0002	100.00
TP	295.2918	93.92	154.4425	96.96	748.2859	78.43	1424.8349	71.62	1018.6807	72.59	162.3835	94.64
T	407.9318	89.44	208.2372	93.72	1225.0221	60.15	1849.5477	64.96	1450.2972	52.79	284.6784	82.65
S	1317.5382	70.04	530.1380	84.82	1290.8198	60.88	2482.2654	49.13	1406.7393	54.31	576.7219	61.20
N	841.2034	80.61	565.2425	84.82	1465.7977	64.17	1904.9076	67.90	1075.4100	73.94	223.7113	92.43
O	809.8129	79.74	256.3962	92.71	1598.1539	53.56	2099.6529	55.79	1607.7885	43.15	391.4764	73.82
P	1621.5088	62.52	953.9071	68.62	1880.2138	42.60	3388.4342	36.79	1634.2959	49.41	749.2408	50.47
Valentine biogeographic scheme												
TSONP	240.0002	100.00	150.0002	100.00	90.0002	100.00	180.0001	100.00	90.0002	100.00	210.0001	100.00
TSON	208.0002	100.00	130.0001	100.00	78.0002	100.00	156.0001	100.00	78.0001	100.00	182.0001	100.00
TSO	187.1492	97.83	129.1102	98.18	42.0002	100.00	85.7955	100.00	42.0002	100.00	104.4690	99.00
TSN	176.0001	100.00	110.0001	100.00	66.0001	100.00	132.0001	100.00	66.0002	100.00	154.0002	100.00
TSP	237.5460	96.67	102.3551	98.99	57.3841	99.63	125.9436	99.02	42.0002	100.00	98.0002	100.00
TON	176.0001	100.00	110.0001	100.00	86.3306	99.27	132.0002	100.00	66.0002	100.00	154.0002	100.00
TOP	184.5819	97.83	124.3652	97.77	82.1000	97.99	125.4996	99.34	42.0002	100.00	123.4235	98.00
TNP	176.0002	100.00	110.0002	100.00	66.0014	100.00	132.0002	100.00	66.0002	100.00	154.0001	100.00
TS	240.4446	95.66	122.8495	97.17	50.0358	99.27	132.7752	98.14	30.7640	100.00	105.0760	98.00
TO	241.0551	95.66	226.3560	92.51	86.0032	97.26	181.7345	97.93	30.0008	100.00	145.3042	93.00
TN	144.0001	100.00	135.6083	98.58	79.7823	98.72	125.3002	99.34	54.0001	100.00	126.0001	100.00
TP	239.1273	95.37	129.4160	96.36	86.7303	97.44	192.1835	97.16	31.5758	100.00	148.2403	96.00
T	373.0334	91.61	253.6596	89.68	122.8313	96.34	308.8324	93.89	47.4162	99.15	207.7020	87.00
S	1588.5481	64.40	558.5582	81.17	297.8086	91.22	1195.0286	76.53	351.1043	91.03	599.7627	64.00
O	680.0072	80.75	297.3890	89.88	121.2182	96.16	461.0094	90.72	55.3802	98.31	303.8774	82.00
N	672.2541	88.28	596.1660	80.57	235.7785	93.24	631.5262	90.17	253.0103	95.26	297.8504	89.00
P	1574.1922	63.10	915.3737	69.84	347.3264	91.41	1252.5992	82.31	44.6224	98.98	797.9237	60.00

Environmental variables include: temperature (T), salinity (S), dissolved oxygen concentration (O), nutrient concentrations (N, including phosphate, nitrate, and silicate), and net primary productivity (P). Both means and seasonal ranges are used for each variable. Percentage correct is the percent of grid cells classified into the correct biogeographic unit. AIC, Akaike information criterion.

Table S8. Mantel correlations between bivalve composition and mean annual values and seasonal ranges in temperature, salinity, and net primary productivity

	Pearson		Spearman	
Region	Mantel *r*	*P* value	Mantel *r*	*P* value
Jaccard dissimilarity				
Northwestern Atlantic	0.4001	0.001	0.5041	0.001
Southwestern Atlantic	0.4577	0.001	0.6502	0.001
Northeastern Atlantic	0.3852	0.001	0.4793	0.001
Southeastern Atlantic	0.4981	0.001	0.4982	0.001
Indian Ocean	0.2559	0.001	0.2864	0.001
Northwestern Pacific	0.432	0.001	0.564	0.001
Southwestern Pacific	0.4319	0.001	0.522	0.001
North West Pacific Islands	0.5622	0.001	0.5767	0.001
South West Pacific Islands	0.6765	0.001	0.7399	0.001
Northeastern Pacific	0.5523	0.001	0.6365	0.001
Southeastern Pacific	0.5491	0.001	0.5538	0.001
Simpson dissimilarity				
Northwestern Atlantic	0.3962	0.001	0.5347	0.001
Southwestern Atlantic	0.434	0.001	0.6645	0.001
Northeastern Atlantic	0.383	0.001	0.4898	0.001
Southeastern Atlantic	0.3141	0.001	0.354	0.001
Indian Ocean	0.3136	0.001	0.3484	0.001
Northwestern Pacific	0.3754	0.001	0.5632	0.001
Southwestern Pacific	0.4446	0.001	0.543	0.001
North West Pacific Islands	0.7651	0.001	0.703	0.001
South West Pacific Islands	0.6982	0.001	0.766	0.001
Northeastern Pacific	0.5165	0.001	0.6205	0.001
Southeastern Pacific	0.5348	0.001	0.5353	0.001

P values were estimated after 1,000 iterations.

Dataset S1. The 1° grid-cell identifiers and environmental data

Dataset S1

Dataset S2. Jaccard-distance matrix of bivalve assemblage data at the 1° grid-cell resolution for the Atlantic Basin

Dataset S2

Dataset S3. Simpson-distance matrix of bivalve assemblage data at the 1° grid-cell resolution for the Atlantic Basin

Dataset S3

Dataset S4. Jaccard-distance matrix of bivalve assemblage data at the 1° grid-cell resolution for the Pacific Basin

Dataset S4

Dataset S5. Simpson-distance matrix bivalve assemblage data at the 1° grid-cell resolution for the Pacific Basin

Dataset S5