

Chapter 2

Techniques

Abstract Databases used for the analysis of past biotas should be as internally consistent as possible taking into account the incompleteness of the fossil record and the taxonomic distortions due to the history of their knowledge. A comprehensive and critically updated database of Southern Hemisphere bivalve occurrences through the Triassic and Jurassic was built. Most of paleobiogeographic analyses were performed within time slices to obtain comparable results in a time succession. Analytical methods were used for both (a) the analysis of latitudinal ranges along the South American paleo-coasts, and (b) the recognition of paleobiogeographic units for the Southern Hemisphere. a) The first approach to the study of species latitudinal ranges was cluster analysis, but this method, although useful, imposes a hierarchical structure on the data. Thus, to check for faunal changes along latitude, the distribution limits of species were explored using a technique similar to that considered for origination/extinction analysis, substituting first and last appearances by northernmost and southernmost geographical occurrences. Generalized linear models were used to look for changes on the proportional values of different species categories related to systematic and paleobiogeographic kinships. b) For the recognition of biochoremas, the incomplete and uneven nature of the data precludes the application of methods which may group areas according to the common absence of data, and we followed a traditional approach based on endemism. In order to check the biogeographic structures without assuming a hierarchical or gradational arrangement, a Bootstrapped Spanning Network was calculated.

2.1 The Data

Occurrences of Triassic and Jurassic bivalve species were compiled from various published sources as well as the authors' own data, and plotted stage by stage from Induan (Early Triassic) to Berriasian (Early Cretaceous). Ros' (2009; also

Ros et al. 2012) updated database was used for Triassic occurrences. The study area is restricted to the Paleo-Southern Hemisphere, but the initial database was compiled on a global scale, not only to provide the necessary framework for the detailed analysis of southern regions, but also to adequately recognize patterns of general distribution and endemism. The species distribution data compilation was systematically and stratigraphically updated as far as possible and dubious records were excluded. The most serious problems related to such global databases are the incompleteness of the fossil record and the taxonomic distortions introduced by different authors working in different areas at different times. The first problem has no immediate solution, while the second can be somewhat reduced by critical evaluation of the data. This cannot always be done, for several reasons, but internal consistency was sought whenever possible, with careful reappraisal of both taxonomy and age of the records taken from the previous literature. In order to obtain a sound foundation for biogeographic considerations, only species personally examined or adequately described and figured were included; uncritical listing of taxa from sources lacking illustrations was avoided.

Presence-absence data were used throughout, since reliable quantitative records are only available for a small fraction of the occurrences.

Theoretically, the species is the most objective of taxonomic units; however, when species lists are compiled from studies made by various authors and at different times they become intensely subjective to the point that compilations at the generic or familial levels are preferred for global analysis (Stehli et al. 1967). Furthermore, the use of genus-group taxa increases the consistency of the database, as generic and sub-generic concepts have more consensus than species among different authors. In this book, genera and subgenera are used for the larger scales analyses (hemispheric and global), while species are preferred for performing regional analysis within the area studied by the authors, where first-hand knowledge facilitates identification and consistency.

Although it is evident that there are serious gaps in our knowledge of Triassic and Jurassic faunas from certain regions and for some bivalve groups, which have not been systematically reviewed or updated yet, it is believed that the data are comprehensive enough for the general purpose of this study.

In order to rationalize the study, the biogeographic affinities of the species (according to the categories discussed in Sect. 3.3) were recorded as well. Since the assignment of species to a definite type is based on its known distribution and relative abundance, and this knowledge is constantly being improved, this task proved difficult in some instances but relatively straightforward in others (such as the pectinaceans, see Hayami 1989; Damborenea 1993). Though many of the species have local distributions restricted to South America, they may have strong relations with other species or genera belonging to the paleobiogeographic affinities categories recognized here, and these were listed and used in the analysis.

Relative abundance was regarded as an important factor too; sporadic occurrences outside the main area of distribution are to be expected and, if adequately recognized as such, should not obscure the picture.

2.2 Quantification: A Difficult Approach

The application of any analytical technique in biogeography implies the acceptance of a defined biogeographic theory; biogeographic patterns are the result of ecological and historical processes acting on them, and if one accepts that such patterns actually do exist, then some analytical approach is required (Lieberman 2000). Quantitative analysis are claimed to be necessary in order to make biogeography a more rigorous science; these kinds of approaches should act as arbiters to choose between competing hypotheses. Nevertheless, these analytical techniques on their own do not guarantee valid results (Lieberman 2000), and the quality of the outcome largely depends of the quality of the available data.

The basic unit of any biogeographic study is the geographic range of a species, but its characterization is complicated by problems of defining and mapping them; what we usually see as maps of geographic ranges are simplifications of complex historic and ecologic patterns (Brown et al. 1996). Species ranges are usually mapped as irregular continuous areas, although in some cases they are presented as more precise “dot maps” that plot each location where a species has been recorded; most published range maps attempt to define the historical range of a species, encompassing localities where a species has regularly occurred in the past and recently colonized areas (Brown et al. 1996). To all these problems, common in the analysis of modern biota, the imperfection of the fossil record and the time averaging are added on paleobiogeographic analyses.

We use reconstructions of Mesozoic paleogeography to discuss our results. As pointed out by Smith (2011), although we now know a great deal more than 50 years ago, available Mesozoic paleogeographic maps are still rudimentary, showing only where continents and oceans were in the past, without bathymetry data, for instance. For the general purposes of this book they are nevertheless adequate.

Like any other model, a biogeographic model is a simplification of reality, and hence, as pointed out above, considered geographic ranges are simplified from actual geographic ranges; the kind of simplification applied will depend on the aim of the study.

In Sect. 4.2, an analysis of bivalve distributions along the west paleo-coast of South America is addressed; this coast presents a north–south orientation nowadays and it was similarly disposed during the Jurassic. This peculiarity allows for an analysis of latitudinal gradients on bivalve distribution. These gradients are associated with climatic gradients, imposing ecological restrictions to species distribution. A total of 13 areas were defined between parallels 20° and 46° S, each with a latitudinal range of 2°, as considered by Damborenea (1996). The presence of a species in one locality was computed as a presence for the whole area. During this analysis, discontinuity of the outcrops and differences in preservation probability were interpreted to be biasing a somewhat continuous distribution (Fig. 2.1), and so species ranges were interpolated between two localities with occurrence

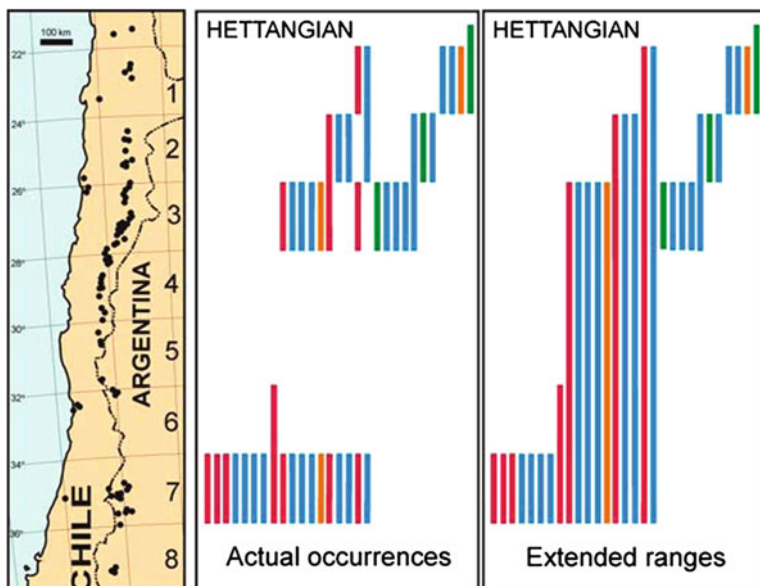


Fig. 2.1 Latitudinal range data exemplified with the Hettangian species data along a section of the eastern Paleo-Pacific coast. The actually observed occurrences (*left*) were used in most of the analyses performed, while the graphic presentation of data in [Sect. 4.2.2](#) (Figs. 4.7, 4.9, 4.11 and 4.13) was simplified to show only the extended ranges (*right*)

(although on certain analyses, explained below, only the observed data were considered).

In [Sect. 5.2](#), a hemispheric-scale analysis was carried out, and for this the units considered were well defined and bounded regions established a priori based on the current literature on the topic. Presence or absence of taxa (genera group taxa in this case) was computed for each region as a whole, disregarding the internal structure of distribution. The detail lost with this approach is not relevant at the scale considered for the purposes of this review.

Concerning time, most of the analyses performed were developed within time slices coinciding with the different stages (or group of stages), as a way of obtaining comparable results in a time succession. If stages are used, time slices are very unevenly “thick” (see [Table 1.2](#), in which the vertical scale is proportional to time in Ma). One way to overcome in part the unevenness of the time duration of each slice is to group stages as done for some of the analyses (see [Table 5.1](#)), a method which also has some other practical advantages during the data gathering process.

On top of this, the time averaging within each of these time slices is considerable. The effects of time averaging are difficult to remove, especially in small-scale studies; as mentioned before, “instantaneous” geographic ranges cannot be obtained even for modern species, and this is yet less likely for fossil species. To deal with this, time bins were considered, and stratigraphic ranges are bounded by

their upper and lower limits. The geographic ranges presented here include all the localities, where a species was present at any time of that time bin, in this case a stage; in some cases, the paucity of data forced the grouping of stages analyzed.

2.3 Analytic Methods

Phenetic approaches are common techniques in biogeographic analysis; in these methods information about distribution of taxa is entered into a data matrix for each of the regions or areas analyzed (Lieberman 2000). For this particular work only presence or absence of taxa (species or genera, depending on the analysis) was considered; abundance was indirectly considered solely when establishing biogeographic affinities.

All calculations were done using the different facilities provided by the softwares PAST (Hammer et al. 2001) and R (R Development Core Team 2008), as well as standard spreadsheet software.

Analytical methods further used to deal with both the latitudinal distribution and the biochoremas recognition are shortly described below.

2.3.1 *Latitudinal Distributions*

2.3.1.1 Cluster Analysis

The first explorative technique applied here is the hierarchical cluster analysis, for which a distance or similarity measure must be defined (Hammer and Harper 2006). Our main goal was to group together the localities according to their species content, so the Simpson's coefficient of similarity (Simpson 1943; see also Shi 1993) was used. This index is defined as the number of shared species between two localities divided by the number of species in the smaller sample. This index is totally insensitive to the size of the larger sample, which makes it suitable when sampling is considered to be incomplete (Shi 1993; Hammer and Harper 2006), as is the case for our database. The localities considered here were not equally treated in the literature, neither have the same abundance of fossils; hence, they cannot be considered as equally sampled, making the Simpson's coefficient the most adequate available index of similarity to use. Cluster analysis is an ordination method, grouping elements according to their similarity; clusters or groups have no statistical significance associated. A support value can be obtained for the nodes by simply resampling taxa (in this case species) and building a new dendrogram; the proportion of times the node appears on the dendrograms resulting from the resampled matrices is the support value for the node. Although the general grouping and disposition of the localities are evaluated on each analysis, special value is given to groups with similarity values of 0.50 or higher (i.e., 50 % of

species shared or more) and to groups with support values of 0.50 or higher, as considered in other paleobiogeographic studies (Brayard et al. 2007; Dera et al. 2011).

On the analysis of latitudinal gradients the main interest focuses on distribution limits, so for species that appear at two distant localities it is usual to extend their ranges along the intermediate latitudes, and most data here are presented graphically in that way for clarity (Fig. 2.1). Nevertheless, for this analysis that methodology would result in a circular reasoning, since the latitudinal gradient would be analyzed presuming its existence; nearby localities would be similar because we assume they share species for being close to each other. To avoid this, cluster analysis was performed on the actually observed presence/absence data; this may produce some sensitivity to differences in knowledge between localities, but that is why we use Simpson's coefficient.

2.3.1.2 Distribution Limits of Species

To check for the faunal changes along a latitudinal gradient, we analyzed the distribution limits of the considered species through that gradient. Cluster analyses, although useful, are hierarchical ordination methods and hence they impose a hierarchical structure on the data, whether this exists or not. If a gradation among localities is to be expected, as happens in a latitudinal gradient, other independent approaches should be considered to check for it. A first graphic and very simple approach is to analyze the distribution limits of the considered species through that gradient. The methodology applied is similar to that considered for origination/extinction analyses, counting the first and last appearance data (FAD and LAD respectively) on each stage (Hammer and Harper 2006), although in this case the stages are substituted by the latitudinal intervals, while the FADs and LADs are replaced by the northern distribution limit datum (NDL) and the southern distribution limit datum (SDL). If faunal turnover presents a gradational pattern, then high values of SDL and NDL are expected in all areas. On the other hand, sudden changes in faunal distribution will be recognized as peaks on the graphic; particularly significant will be the coincidence of peaks on both curves since they will show a major faunal turnover at that latitude (i.e., there will be a lot of species that appear only to the north and a lot that appear only to the south of that point). Peaks on only one curve indicate a reduction on general diversity on one direction (either north or south) and may be informative depending on the nature of data. This reduction could be spurious, if it only represents a sampling bias. For example, on the graphic for the Pliensbachian stage (Fig. 4.14), there is a high peak on the NDL curve between 24° and 26° S, but data for the areas between 20° and 24° S are scarce, and hence many of the considered species may have a broader range, extending northwards; something similar may be happening on the SDL peak between 40° and 42° S.

2.3.1.3 Generalized Linear Models

A second approach to check for gradational patterns is to look for changes on the proportional values of different species categories; data such as biogeographic affinities or systematic kinship (for instance, superfamilies) are good raw material for this kind of analysis. Generalized linear models (GLMs) are useful for data on proportions (Crawley 2007). The software R (R Development Core Team 2008) carries out a weighted regression, using the individual sample sizes as weights and the logit link function to ensure linearity (Crawley 2007). As a result a linear predictor is obtained together with its significance; the significance level used here was 0.05, but significance values between 0.05 and 0.10 were also considered as potentially explanatory. Positive linear predictors will imply positive associations between variables (i.e., an increment in the independent value, on this case latitude, is associated to an increment on the dependent value, on this case the proportion of species of the analyzed group). Negative linear predictors will imply the opposite trend, i.e., an increasing proportion of species of the group toward the north. This same analysis was applied to other groupings, such as superfamilies; on this last case the analyses were formed both on observed data and on extended range data. Localities poorly sampled may introduce noise instead to clear up a pattern, and so were removed from the analysis. When this happened it was made clear in the discussion.

2.3.2 Recognition of Biochoremas and Their Characterization

The recognition and characterization of “conventional” biogeographic units, i.e., based on area and hierarchy, was dismissed by some authors because “objective” criteria are lacking and methods used are not rigorous enough. One of the aims of biogeography is to infer historical connections among biotas or geographical areas on the basis of the distribution of organisms (historic biogeography). A variety of analytical methods were developed to compare the biota of different localities with this purpose (see a discussion and references in Posadas et al. 2006), one of the most popular being parsimony analysis of endemism (PAE) developed by Rosen (1988).

The incomplete and uneven nature of our data precludes the application of methods which may group areas on the basis of the common absence of taxa. So at this stage we have mostly followed a traditional qualitative approach based on endemism, combined with other features, such as diversity, to characterize biochoremas. In the same vein, we have consciously avoided the use of ranking names for the units so discussed (such as Realm, Province, etc.), and prefer to refer them simply as “biochoremas”, sometimes specifying their relative ranking only.

Cluster analysis was applied to these data too, but only as a first approach since, as explained before, this kind of analysis imposes a hierarchical structure.

2.3.2.1 Bootstrapped Spanning Network

In order to check biogeographic structures between regions but without assuming either a hierarchical or a gradational structure, we calculated a Bootstrapped Spanning Network or BSN (Brayard et al. 2007) for the localities involved. The construction of a BSN is a straightforward procedure; starting from a matrix of observed occurrences a dissimilarity matrix was obtained, using as distance measure the complement value of Simpson's coefficient (1-S). From this second matrix a minimum spanning network (MSN) was obtained, by connecting localities from lower to higher dissimilarity value until all localities are directly or indirectly connected. Then, the original data matrix of observed occurrences was randomly resampled with replacement, and new dissimilarity matrix and MSN were calculated; this whole procedure was repeated 1000 times. For each edge on the original MSN a bootstrapped support value was obtained by simply calculating the proportion of times each edge was obtained on the original and resampled MSNs. Starting from the observed MSN, edges were removed (starting by those with the least support value) if and only if the network was still connected after its removal and the overall product of bootstrapped supports was increased; the resulting network is the bootstrapped spanning network (Brayard et al. 2007), obtained here for the more complete data sets.

BSN may be sensitive to the presence of sampled areas with low diversity. Lets consider as an extreme case a locality with only one taxon; if this taxon is endemic to that region, then the distance value of this area will be maximum (i.e. 1) with any other area. During the construction of the resampled MSN, if this taxon is sampled, all values lower than 1 (almost all edges) will be chosen, hence incrementing the support values of many edges that most probably should be considered as low supported. To avoid this, areas with less than 10 known taxa for the stage involved were removed from the analysis.

References

- Brayard A, Escarguel G, Bucher H (2007) The biogeography of Early Triassic ammonoid faunas: clusters, gradients and networks. *Geobios* 40:749–765
- Brown JH, Stevens GC, Kaufman DM (1996) The geographic range: size, shape, boundaries, and internal structure. *Ann Rev Ecol Syst* 27:597–623
- Crawley MJ (2007) *The R book*. Wiley, NY
- Damborenea SE (1993) Early Jurassic South American pectinaceans and circum-Pacific palaeobiogeography. *Palaeogeogr Palaeoclimatol Palaeoecol* 100:109–123
- Damborenea SE (1996) Palaeobiogeography of Early Jurassic bivalves along the southeastern Pacific margin. 13° Congr Geol Argent, 3° Congr Explor Hidrocarb (Buenos Aires). *Actas* 5:151–167
- Dera G, Neige P, Dommergues JL, Brayard A (2011) Ammonite paleobiogeography during the Pliensbachian–Toarcian crisis (Early Jurassic) reflecting paleoclimate, eustasy, and extinctions. *Global Planet Change* 78:92–105
- Hammer Ø, Harper DAT (2006) *Paleontological data analysis*. Blackwell Publishing, Oxford

- Hammer Ø, Harper DAT, Ryan PD (2001) PAST: Paleontological statistics software package for education and data analysis. *Palaeontol Electron* 4(1):9
- Hayami I (1989) Outlook of the post-Paleozoic historical biogeography of Pectinids in the Western Pacific Region, The Univ Mus, Univ Tokyo. *Nature Culture* 1:3–25
- Lieberman BS (2000) Paleobiogeography. Using Fossils to study global change, plate tectonics, and evolution. *Topics in Geobiology* 16, Kluwer Academic, Plenum Publishers, New York
- Posadas P, Crisci JV, Katinas L (2006) Historical biogeography: a review of its basic concepts and critical issues. *J Arid Environ* 66:389–403
- R Development Core Team (2008) R: a language and environment for statistical computing [Internet]. Vienna: R Foundation for Statistical Computing. <http://www.R-project.org>
- Ros S (2009) Dinámica de la paleodiversidad de los Bivalvos del Triásico y Jurásico Inferior. PhD Thesis. Univ Valencia. Valencia. <http://www.tesisenred.net/handle/10803/9952>
- Ros S, Márquez-Aliaga A, Damborenea SE (2012) Comprehensive database on Induan (Lower Triassic) to Sinemurian (Lower Jurassic) marine bivalve genera and their paleobiogeographic record. *Paleontol Contrib, Univ Kansas* (in press)
- Rosen BR (1988) From fossils to earth history: applied historical biogeography. In: Myers AA, Giller PS (eds) *Analytical biogeography: an integrated approach to the study of animal and plant distributions*. Chapman and Hall, London
- Shi GR (1993) Multivariate data analysis in palaeoecology and palaeobiogeography—A review. *Palaeogeogr Palaeoclimatol Palaeocol* 105:199–234
- Simpson GG (1943) Mammals and the nature of continents. *Am J Sci* 241:1–31
- Smith AG (2011) Uncertainties in Phanerozoic global continental reconstructions and their biogeographical implications. In: Upchurch P, McGowan AJ, Slater SC (eds) *Palaeogeography and Palaeobiogeography. Biodiversity in Space and Time. Syst Assoc Spec* 77:39–74
- Stehli FG, McAlester AL, Helsley CE (1967) Taxonomic diversity in recent bivalves and some implications for geology. *Geol Soc Am Bull* 78:455–466

Southern Hemisphere Palaeobiogeography of
Triassic-Jurassic Marine Bivalves

Damborenea, S.E.; Echevarría, J.; Ros-Franch, S.

2013, VIII, 139 p. 47 illus., 34 illus. in color., Softcover

ISBN: 978-94-007-5097-5