1. Introduction

The geographic distribution of organisms is the subject of Biogeography, a field of biology that naturalists have carried out for over two centuries [1-6]. From the observation of animal and plant distribution, diverse questions emerge; the description of diversity gradients; delimitation of areas of endemism; identification of ancestral areas and search of relationships among areas, among others, have become major issues to be analyzed, worked out and solved. In this way, biogeography has turned into a multi-layered discipline with both theoretical and analytical frameworks and far-reaching objectives.

However, at the beginning it was closely related to systematics. Taxonomists were the ones who took a keen interest in the geographical distribution of taxa. In other words, because the connection is so close, several analytical tools applied to the treatment of biogeographical problems are adaptations or modifications from methods oriented to solve systematics questions. This apparent panacea may also represent one important analytical obstacle for biogeography. Although some biogeographical questions require systematic information to be solved, the object of study of biogeography, that is, spatial distribution of taxa, as well as its concepts and problems, are different from those of systematics. Hence, methods taken from systematics are not appropriate for the treatment of biogeographical problems. The need for its own methods and its own analytical framework have promoted prolific theoretical discussions and methodological developments throughout the last 20 years. In this context, the concept of areas of endemism is being widely debated and several methods have been proposed to attempt to identify these patterns. Areas of endemism have a central role in biogeography as they are the analytical units in historical biogeography, and are also considered quite relevant for biodiversity conservation [7]. It is the aim of this chapter to introduce
the major discussions around the concept of areas of endemism and focus on analytical problems associated with its identification. A brief revision of contributions on endemism in South America is presented and some limitations associated to empirical analysis are highlighted in order to give an overall picture on the current state of affairs on this controversial subject.

2. Areas of endemism, its importance

In biogeography, the term “area of endemism" is used to refer to a particular pattern of distribution delimited by the distribution congruence of, at least, two taxa [8]. Given that the range of distribution of a taxon is determined by historical, as well as current factors, it can be assumed that those taxa which show similar ranges have been affected by the same factors in a similar way [9]. The identification of areas of endemism is an essential first step to elaborate hypotheses that help to disclose the general history of biota and the places where they inhabit. Because of this, recognition of these patterns has been central to biogeography. Oddly enough, and despite its indisputable importance, endemism involves several problems which reach even its definition (semantic field), not to mention those resulting from the absence of a clear framework (conceptual problem) or those associated to the identification of areas of endemism (analytical issues) [9, 11-19]; While the first two problems are briefly dealt with in the present chapter, identifying and assessing the main areas of endemism will be the main focus.

2.1. Defining the term

The idea of endemism dates back to more than 200 years, and has been employed, as it is actually understood, by de Candolle [1]). Since then, the concepts of endemicity and areas of endemism have been widely discussed. Some problems around these concepts emerge from the diverse uses and interpretations given to them in literature (e.g. [16, 20-21]. Harold and Moii [21]. Although differences between diverse uses as regards connotations could seem minor, the lack of precision in the definition of these concepts hinders an unambiguous interpretation and causes confusion. Additionally, numerous expressions, such as "generalized track”, “track”, “biotic element”, “centers of endemicity”, “units of co-occurrence”, among others, are commonly used as synonyms of area of endemism, [16, 21-23]. Although basically related with the term "areas of endemism”, these concepts refer to different patterns of distribution and are defined on different theoretical grounds.

3. A clear conceptual framework

As it usually happens in other fields such as morphology and embryology, in the field of biogeography, the identification and description of patterns precede the inference of the causes of its occurrence. However, some biogeographers assume that vicariance must be involved [12, 17]). According to this idea, a pattern of sympatry among species could be defined as area of
endemism only if it emerged from a vicariant event. This assumption entails new difficulties for the identification of areas of endemism: the causes which originate the patterns must be known a priori, or else, the identification of patterns and processes should be performed simultaneously. Fortunately, most biogeographers follow the generalized concept, which supposes that multiple factors affect and define current patterns.

4. Identifying areas of endemism

The identification of such areas has been a major challenge in biogeography and deals with several difficulties, some of them related with the two questions mentioned above. However, in the last decades, several methods for identification of these patterns have been proposed [9, 15-16, 18, 24-25] In general, current methods for recognizing areas of endemism can be classified on the basis of whether they aim to determine (i) species patterns, i.e. groups of species with overlapping distributions, or (ii) geographical patterns, i.e. groups of area units with similar species composition. These approaches assess closely related but slightly different aspects of biogeographical data. Methods dealing with species patterns group species with similar distributions and result in clusters -which may or may not define obvious spatial patterns-. Instead, methods oriented to define geographical patterns, are more related to the classical notion of area of endemism, resulting in geographical areas defined by species distributions.

The methods currently in use are many and heterogeneous. While reflecting the multiple conceptions of areas of endemism, these proposals differ in their theoretical bases as well in its mathematical formulations. Following are three of them: Parsimony Analysis of Endemicity (PAE [15]), Biotic Elements (BE; Hausdorf and Hennig, 2003[24]), and Endemicity Analysis (EA; Szumik et al., 2002[9]; Szumik and Goloboff, 2004[18]). Although several modifications of PAE, as well as other hierarchical methods have been proposed (see [16, 26]), this method has been selected as a representative of hierarchical methods because it remains the most widely used in empirical analyses ([27-33].

**PAE.** The Parsimony Analysis of Endemicity (PAE) was the first method proposed to formally identify areas of endemism[15]. The input data for PAE consist of a binary matrix in which the presence of a given species (rows) in an area unit (columns) is coded as 1 and its absence as 0. Analogous to a cladistic analysis, PAE hierarchically groups area units (analogous to taxa) based on their shared species (analogous to characters) according to the maximum-parsimony criterion. Therefore, PAE attempts to minimize both “dispersion events” (parallelisms) and “extinctions” (secondary reversions) of species within a given area. Areas of endemism are defined from the most-parsimonious tree (or strict consensus) as groups of area units supported by two or more “synapomorphic species” (i.e. endemic species [15]). In its most classical formulation, species that present reversions (i.e. are absent in any of the area units) and/or parallelisms (i.e. are present elsewhere) in their distributions are not considered endemic. Therefore, PAE is especially strict when penalizing the absence of a species within an area, which makes it more likely to fail to detect a relatively large number of areas of endemism.
Despite the well-known limitations of hierarchical classification models in the delimitation of areas of endemism [9, 33-34]), PAE remains the most widely used method for describing biogeographical patterns [31- 32, 35]).

**BE.** Hausdorf [17] considers areas of endemism in the context of the vicariance model, and argues for the use of “biotic elements” defined as “groups of taxa whose ranges are significantly more similar to each other than to those of taxa of other such groups” (p. 651[17]), rather than the more traditional areas of endemism [24]). This method is implemented in the R package Prabclus by Hennig [36], which calculates a Kulczynski dissimilarity matrix [37]) between pairs of species which is then reduced using a nonmetric multidimensional scaling (NMDS; [38]). A Model-Based Gaussian clustering (MBGC) is applied to this matrix to identify clusters of species with similar distributions, or biotic elements. In spatial terms, a biotic element is equivalent to the spatial extent of the distributions of all species included in the cluster.

**EA.** In 2002, Szumik and colleagues proposed an optimality criterion to identify areas of endemism by explicitly assessing the congruence among species distributions. This proposal, improved by Szumik & Goloboff [17]), is implemented in NDM⁄VNDM by Goloboff [39] and Szumik and Goloboff [9]). The congruence between a species distribution and a given area is measured by an Endemicity Index (EI) ranging from 0 to 1. The EI is 1 for species that are uniformly distributed in the area under study, and only within that area (“perfect endemism”), and decreases for species that are present elsewhere, and / or poorly distributed within the area. In turn, the endemicity value of an area (EIA) is calculated as the sum of the EIs of the endemic species included in the area. Therefore, two factors contribute to the EIA: the number of species included in the area and the degree of congruence (measured by the EI) between the species distributions and the area itself (for details see [9]).

The emergency of quantitative methods that allow describing these patterns objectively has represented an important advance in the discussion of endemism. However, the contrast between different methodological proposals introduced new questions: are the hypothesis resulting from different analysis homologous? Is there a better method to identify areas of endemism? A few recent contributions attempt to elucidate these queries by testing and exploring the behaviour of some methods, e.g., [34, 40]. Several comparisons between methods have been performed by using real data [41-43]). However, real data provide only a limited assessment of the differences between the procedures. Some characteristics of the distribution of species, e.g. geographical shape or number of records, affect pattern recognition in uncertain ways. Furthermore, sampling bias, which often affects available distributional data, causes problems in the identification of biogeographical patterns [44]). As it is often difficult to distinguish whether the identified patterns result from singularities of the data or properties of the methods, an evaluation based on real datasets, or data simulated under realistic conditions, is not enough to establish general conclusions on the performance of the methods.

Recently, Casagranda et al.[19]) states a comparison by using controlled -hypothetical distributions, pointing differences, advantages and limitations of Endemicity Analysis (EA), Parsimony Analysis of Endemicity (PAE), and Biotic Elements Analysis (BE) in their study, these authors measured the efficiency of the methods according their ability to identify
hypothetical predefined patterns. These patterns represent nested, overlapping, and disjoint areas of endemism supported by species with different degrees of sympatry.

This comparison shows how the application of different analytical methods can lead to identification of different areas of endemism, and reveals some undesirable effects produced by methodological idiosyncrasies in the description of these patterns. Following are the main results reported in this contribution:

PAE shows a poor performance at identifying overlapping and disjoint patterns. In all cases, PAE is able to recover areas defined by perfectly sympatric species, but its performance decreases as the incongruence among the species distributions increases (Figure 1).

As regards BE, it is very sensitive to the degree of congruence among the distributions of the species that define an area, showing a counterintuitive behaviour: while the method cannot recognize patterns defined by perfectly sympatric species, its performance improves with increasing levels of incongruence between the species distributions. BE often report multiple distinct biotic elements for species which actually have very similar distributions (Figure 2 a) as well as reporting a single biotic element including species with completely allopatric distributions (Figure 2 b). These examples show discordance between the theoretical basis of
the approach [16]) and its practical implementation. Together, these limitations suggest the
users should exercise caution when interpreting the results generated by this method.

![Graph showing the number of recovered areas by subcase for different cases. (a) Case 0: mixed noise, adjacent noise, internal noise, no noise. (b) Case 1: mixed noise, adjacent noise, internal noise, no noise. (c) Case 2: mixed noise, adjacent noise, internal noise, no noise. (d) Case 3: mixed noise, adjacent noise, internal noise, no noise.]

Figure 2. Special results found by biotic elements. (a) Three species with similar distributions (sp.a, sp.b, and sp.c) are separated in different biotic elements (BE 1, BE 2 and BE 3); (b) three species with completely allopatric distributions (sp.d, sp.e, and sp.f) are grouped in the same biotic element (BE 4) (Modified from Casagranda et al., 2012.).

Regarding EA, it shows a high percentage of success in the recovery of predefined areas with no discrimination of case, whether nested, overlapping or disjoint, of degree of congruence between distributions of species. EA reports frequently redundant “twin” areas that have only slight differences in spatial structure and/or in their species composition.

Taking into account that overlapping and disjoint patterns are relatively common in nature, and that, in general, sympathy between species varies widely, PAE is probably not the most suitable method to describe areas of endemism based on real distributional data. Although ideal cases are not frequently observed on the spatial scale used for most biogeographical analyses, the inability of BE to identify a perfect case of the pattern which the method intends to describe is questionable. The flexibility to recognize areas displayed by EA is associated with the fact that, in contrast to the other methods considered here, EA uses both the number of species and the overlap between their distributions as optimality criteria to search for areas of endemism.

One serious problem is that the method relies on an algorithm that is ineffective for its intended purpose. PAE, for example, is a hierarchical method implying that each cell is included in at least one area of endemism; consequently, PAE cannot describe overlapping patterns, such as nested areas. Additionally, the maximum parsimony criterion aims to minimize the number of homoplasies, resulting in PAE hardly identifying any disjoint areas.

Similarly, BE model-based inference requires a series of distributional assumptions which, if not satisfied, may lead to unreliable or erroneous conclusions. Thus, even if, in theory, a biotic element is defined as a “group of taxa whose ranges are significantly more similar to each
other than to those of taxa of other such groups”, the method may both group totally allopatric species and fail to recognize biotic elements defined by totally sympatric species (see Fig. 2).

An inescapable consequence of the application of an optimality criterion is that multiple hypotheses may be obtained in an analysis; in the case of EA, the “twin” areas represent small variations of single cells. The ambiguity in the input data often results in multiple “best” solutions according to an optimality criterion. The reported alternative and equally optimal patterns often force the researcher to more conservative interpretations.

Conclusions of Casagranda et al. show that EA, in conjunction with consensus areas, is the best available option for endemicity analyses, despite other studies indicating that EA is rather sensitive to certain aspects of the data, such as spatial gaps of information [34]. The advantages of EA over other methods are related to considering spatial information during the identification of areas, as well as using the classical definition of area of endemism as the basis for the analysis: [an area of endemism]... is identified by the congruent distributional boundaries of two or more species, where congruent does not demand complete agreement on those limits at all possible scales of mapping, but relatively extensive sympatry is a prerequisite [8].

5. Areas of endemism in South America

The knowledge about the distribution of species, as well as the geographical patterns, constitute crucial information for biodiversity conservation [7]. Because of this, the study of both species distributions and the mechanisms that give them rise have increased since the awareness of biodiversity crisis.

In the last few years, endemicity has acquired importance in conservation biology since it is considered an outstanding factor for delimitation of conservation areas [45-47]).

Due to its particular history and its huge biodiversity, South America is interesting from a biogeographical point of view. Numerous contributions have been made to address diverse aspects of the distribution of South America’s biota ([47], [48-49] [50-55]; however, quantitative studies are relatively recent.

The development of computational methods [8, 14, 17 23, 35] together with the availability of biodiversity data-bases, such as CONABIO[57] GBIF, [58] y SNDB [59], and Jetz contribution [60] has promoted the advance of empirical analyses dealing with the description of areas of endemism. It is reflected in numerous publications focused on different methodological perspectives and including diverse taxa, in various places of South America [33, 40, 61-64]. A remarkable example of these studies is the recent contribution of Szumik et al (2012) [63], framed between parallels 21 and 32 S and meridians 70 and 53 W, (Figure 3) in the North region of Argentina.

Although the idea of an area of endemism implies that different groups of plants and animals should have largely coincident distributions, most studies of this type are focused on analyzing a restricted number of taxa. In this sense, the analysis of Szumik et al. (2012) represents an
atypical example because the number and diversity of taxa included, more than 800 species of mammals, amphibians, reptiles, birds, insects and plants, representing one of the first approximations to the analysis of total evidence in a biogeographical context.

The quality and structure of data influence the identification of biogeographical patterns [19, 43]. Since the knowledge about distribution of organisms is scarce and taxonomical misidentification and georreferencing errors are commonly observed in available distributional data, an appropriate revision and correction of input information is essential to perform reliable biogeographical descriptions. In this sense, the above mentioned analysis differs from similar studies because the traits of the analyzed data set: “unique among biogeographical studies not only for the number and diversity of plant and animal taxa, but also because it was compiled, edited, and corroborated by 25 practising taxonomists, whose work specializes in the study region Thus, it differs substantially from data sets constructed by downloading data from biodiversity websites” (Szumik et al 2012, p.2[63]; see Figure 3).).

Figure 3. Maps of Argentina: a) relief map; b) biogeographical divisions of Argentina according to Cabrera and Willink (1973); the study region is framed in the red square.

The results reported by these authors indicate that when all the evidence is analysed for a given region, it is possible to obtain areas supported by diverse taxonomic groups (Navarro et al., 2009[63]): half of 126 found areas are supported by three or more major groups. Examples of areas of endemism defined by multiple taxa are the Atlantic Forest (Selva Paranaense—Neotropical, Figure 4) and the north Yungas forest sector (tropical Bermejo- Toldo-Calilegua, two of the most diverse ecorregions of the region.
The patterns of distribution recognized here depict almost all the main biogeographical units proposed in previous studies [26, 47, 49, 51, 53, 54, 55, 60] the Atlantic Forest the Campos (Grasslands) District, the Chaco shrubland (Fig. 5a), the deciduous tropical Yungas forest the Puna highland, and the tropical tails entering Argentina in two disjoint patches[63]. Each of these tropical tails represents part of a broader area that extends towards the north of the South American subcontinent.

Additionally, the species that support the various areas are consistent in general with previous biogeographical studies based on individual groups (plants [32]; snakes [66]; mammals:[66]; insects: [63]; birds: [65,67]), and should be noted that several of these species are currently on red lists of threatened species [68-73]).

Figure 4. An example of an area of endemism identified under differents grids sides (results of Szumik et al., 2012)

6. Final comments

The necessity of quantitative methods that allow a formal description of nature on the basis of available evidence has been an important subject in modern biology. In the last 30 years, both the advances in the field of informatics and the development of computational methods to explore diverse biological questions have been remarkable [74-76].

Biogeography is not foreign to these important advances. When having to compare and evaluate alternative biogeographical hypotheses, biogeographers hold no doubts over the importance of quantitative methods. However, unlike other research areas such as systematics, the richness of biogeography is quite noticeable as far as the number and variety of methodo-
logical proposals are concerned in the attempt to solve a given biogeographical problem. In contrast, those studies where the capacity to explain differences between methods or the quality of the results are put to the test are scarce, as well as anecdotal. The case referred to in the present chapter on the identification of areas of endemism clearly demonstrates the urge of serious and critical studies on biogeography. The formal recognition of areas of endemism is a complex issue; quite a lot has been done in the last few years in order to understand it, but there is still a lot to be done. In addition, the current impending threat on biological diversity urges for methodological improvements conducive to more realistic descriptions of biogeographical patterns.

Acknowledgements

We thank authors of references, specially our colleagues of INSUE. Helpful comments, constructive criticism and generosity from Claudia Szumik are greatly appreciated. Luisa Montivero helped with the English text and Andres Grosso with illustrations. This work was supported by grant PIP-Conicet Nº 1112-200801-00696

Author details

Dra Dolores Casagranda¹² and Dra Mercedes Lizarralde de Grosso²³

1 Instituto de Herpetología, Fundación Miguel Lillo, Tucumán, Argentina

2 Consejo Nacional de Investigaciones Científicas y Técnicas, Tucumán, Argentina

3 Instituto Superior de Entomología (INSUE)-Universidad Nacional de Tucumán, Tucumán, Argentina

References


[16] Linder H P. On areas of endemism, with an example from the African Restionaceae. Systematic Biology 2001; 50 892-912. DOI: 10.1080/106351501753462867


[27] Cracraft, 1991; Patterns of diversification within continental biotas: hierarchical-Congruence among the areas of endemism of Australian vertebrates. Australian Systematic Botany 1991; 4 211-427. DOI: 10.1071/SB9910211


[54] Cei J M. Monographie IV: Reptiles del Centro, Centro-Oeste y Sur de la Argentina. Herpetofauna de las Zonas Áridas y Semiáridas. Museo Regionale di Scienze Naturali Torino 1986; Monogr. 4 1- 527


Morrone J J. A new regional biogeography of the Amazonian subregion, based mainly on animal taxa. Anales del Instituto de Biología de Universidad Autónoma de México, Zoología 2000; 71 99-123.


Straube F C, Di Giacomo A. A avifauna das regiões subtropical e temperada do Neotrópico: desafios biogeográficos. Ciência e Ambiente 2007; 35:137-166


Díaz G B, Ojeda R A. Libro rojo de mamíferos amenazados de la Argentina. SAREM (Sociedad Argentina para el Estudio de los Mamíferos), Mendoza, Argentina. 2000.


