

Biogeographical network analysis of plant species distribution in the Mediterranean region

Maxime Lenormand,^{1,*} Guillaume Papuga,^{2,3} Olivier Argagnon,²
Guilhem De Barros,² Samuel Alleaume,¹ and Sandra Luque¹

¹*Irstea, UMR TETIS, 500 rue JF Breton, FR-34093 Montpellier, France*

²*Conservatoire botanique national méditerranéen de Porquerolles, Parc scientifique Agropolis, 2214 boulevard de la Lironde, 34980 Montferrier sur Lez, France*

³*UMR 5175 CEFÉ, CNRS, 1919 route de Mende, 34293 Montpellier cedex 5, France*

ABSTRACT

Aim. We performed a network analysis of the spatial distribution patterns of plant species using a network approach in order to analyse the biogeographical structure of the French Mediterranean flora.

Location. South of France (Languedoc-Roussillon and Provence-Alpes-Côte d’Azur).

Methods. We used a network approach to identify and characterize biogeographical regions in southern France, based on a large database consisting in millions of geolocalized plant samples corresponding to more than 3,500 plant species. This methodology is performed following five steps, from the biogeographical bipartite network construction, to the identification of biogeographical regions under the form of spatial network communities, the analysis of their interactions and the identification of clusters of plant species based on the species contribution to the biogeographical regions.

Results. The French Mediterranean area is divided into eight statistically significant biogeographical regions that are highly connected spatially. Mediterranean and temperate sub-networks can be distinguished.

Main conclusions. The proposed network approach provides a characterization of the biogeographical structure which confirms some previously recognized floristic regions in southern France, while providing more precise insights into the relationships between them. This approach sheds light on the ecological mechanisms shaping the distribution of Mediterranean biota, and exemplify why fragmented distributions are common in the Mediterranean region.

Keywords. Spatial network, species network, biogeographical regions, plant species, Mediterranean France.

INTRODUCTION

The delimitation of biogeographical regions or bioregions based on the analysis of their biota has been a founding theme in biogeography, from the pioneer work of Wallace (Wallace, 1876), Murray (Murray, 1866) or Wahlenberg (Wahlenberg, 1812) to the most recent advances of (Cheruvilil *et al.*, 2017; Ficetola *et al.*, 2017). Indeed, describing spatial patterns of biodiversity has appeared fundamental to understand the historical diversification of biota, and gain a better understanding of ecological factors that imprint spatial patterns of biodiversity (Graham & Hijmans, 2006; Ricklefs, 2004). Additionally, it has become a key element in the identification of spatial conservation strategies (Funk *et al.*, 2002; Mikolajczak *et al.*, 2015; Rushton *et al.*, 2004). To divide a given territory into meaningful and coherent biogeographical regions, the overall aim is to minimize the heterogeneity in taxonomic composition within regions, while maxi-

mizing differences between them (Kreft & Jetz, 2010; Stoddart, 1992). Although such delineation of biogeographical regions has been based for a long time on expert knowledge of qualitative data collection the increasing availability of species-level distribution data and recent technological advances have allowed for the development of more rigorous frameworks (Kreft & Jetz, 2010). Multivariate methods, such as hierarchical clustering algorithms, have thus been successfully applied in a wide range of studies, from the spatial clustering of Europe on the basis of fauna and flora distributional patterns (Heikinheimo *et al.*, 2007; Mateo *et al.*, 2013; Moreno Saiz *et al.*, 2013; Rueda *et al.*, 2010) and the biogeographical regionalization for plants and vertebrates in sub-Saharan Africa (Linder *et al.*, 2012) to the delineation of zoogeographical and phytogeographical regions in China (He *et al.*, 2016; Zhang *et al.*, 2016) to name but a few. The production of detailed cartographic outputs portraying the differentiation of vegetation into distinct homogeneous biogeographical regions remains difficult, especially where spatial heterogeneity of assemblages is associated with complex environmental gradients (Mikolajczak *et al.*, 2015).

* Corresponding author: maxime.lenormand@irstea.fr

Some regions of the world oppose inherent difficulties due to their highly diversified biota, reflecting complex eco-evolutionary processes. The Mediterranean basin is one of the largest and most important biodiversity hotspots in the world (Blondel *et al.*, 2010; Myers *et al.*, 2000). This region hosts about 25,000 plant species representing 10% of the world's total floristic richness concentrated on only 1% of the world's surface (Greuter, 1991). Additionally, a high level of narrow endemism is a major feature of this species richness (Thompson, 2005). Endemism and richness result in a very heterogeneous region, whose comprehension of spatial patterns of plant distribution is clue to get better insights into past and actual processes shaping biodiversity (Quézel, 1999). The onset of a true Mediterranean climate during the Pliocene and the diverse glacial periods of the Pleistocene (Quézel & Médail, 2004) have shaped the most important phases of plant evolution since the Tertiary (Thompson, 2005). Additionally, due to a long history of human presence, it is worth mentioning that contemporary flora has been widely influenced by human mediated dispersal, land-use and other pressures (Dahlin *et al.*, 2014; Fenu *et al.*, 2014). The French Mediterranean area stretches from the Pyrenees in the south-west to the slopes of the Maritime and Ligurian Alps in the east. It encompasses three zones highlighted as glacial refugia (Médail & Didema, 2009), and the eastern sector represents one of the ten main biodiversity hotspots in the Mediterranean area (Médail & Quézel, 1997). This area represents the northern limit of the Mediterranean climate in the western basin, and thus constitutes a climatic transition from a Mediterranean zone that has a summer drought to a temperate zone less prone to summer drought (Walter & Breckle, 1991 1994). On a finer scale, the climate is more complex with several subtypes and intricated boundaries (Joly *et al.*, 2010; Tassin, 2017).

In order to depict spatial structure in such a complex regional flora, a large dataset is required. While the level of diversity and complexity of such dataset may appear overwhelming at first glance, we have recently witnessed the emergence of network-based approaches for identifying and delimiting biogeographical regions where the presence-absence matrix is represented by a bipartite network. For example, Kougioumoutzis *et al.* (2014) applied the NetCarto algorithm (Guimerà & Nunes Amaral, 2005) in order to identify biogeographical modules within the phytogeographical area of the Cyclades. Similarly, Vilhena & Antonelli (2015) proposed a network approach for identifying and delimiting biogeographical region based on the InfoMap algorithm (Rosvall & Bergstrom, 2008). These methods have the great advantage to be non-parametric in the sense that the number of communities does not need to be fixed *a priori*. In some cases, they can even be applied directly on the biogeographical bipartite network thus avoiding choosing a metric to measure the similarity of species composition between geographical locations

(Vilhena & Antonelli, 2015).

The French Mediterranean territory represents an interesting study area to test such methods, given the excellent knowledge of the spatial distribution of the plant species revealed by botanical inventories (Tison & Foucault, 2014; Tison *et al.*, 2014) and the detailed databases compiled by the French National Botanic Conservatory of Porquerolles and the Alpine National Botanic Conservatory. These organisms have organized the synthesis of historical studies alongside with vast fieldwork campaign that has produced valuable knowledge of species taxonomy and distribution throughout the Mediterranean territory.

In this study, we present a biogeographical network analysis of plant species distribution in the French Mediterranean area at different scales. The objective here is to delineate biogeographical regions, identify groups of species and analyse the relationships between the two entities.

MATERIALS AND METHODS

Dataset and study area

The study area, situated in southern France, encompasses the former Languedoc-Roussillon region (five departments of the current Occitanie region: Pyrénées-Orientales, Aude, Hérault, Gard and Lozère) and the whole Provence-Alpes-Côte d'Azur region. It extends around the entire Mediterranean coastline of mainland France and inland, comprising almost all the Mediterranean hinterland, totalling 558,776 km² (Figure 1). The topography is structured by three major mountain ranges, the Pyrenees in the southwest, the Massif central in the north-west and the Maritime Alps in the north-east. In-between, the landscape is mostly hilly with some lowlands around rivers that flow into lagoons or marshy deltas such as the Camargue. The Rhône is the main structuring river and delimitates western and eastern subregions. Acidic substrates and silicate soils are mainly found in the aforementioned major mountain ranges and in the smaller Maure-Estérel range in southern Provence. The remaining part of the territory is dominated by calcareous or marly substrates (principally Cretaceous and Jurassic), with some significant alluvial zones and small volcanic areas.

The SILENE database¹, has been created in 2006, and is the reference botanical database in the former Languedoc-Roussillon region as well as in Provence-Alpes-Côte d'Azur. It contains historical data gathered from the scientific literature and herbaria along with more recent data coming from public studies, partnerships, local amateur botanist networks and professional botanists of the Conservatoires

¹ Système d'Information sur Les Espèces Natives et Envahissantes - <http://flore.silene.eu/index.php?cont=accueil>

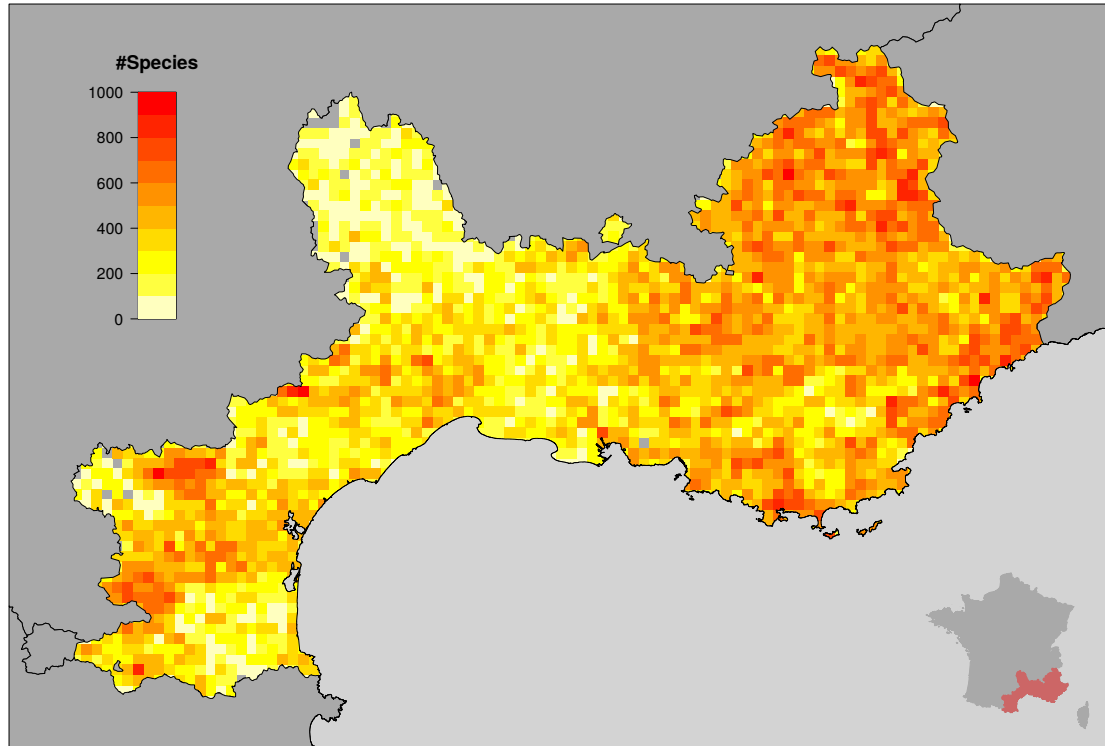


Figure 1. Distribution of the number of species per grid cell ($l=5$ km). The inset shows a map of France including the studied area colored in red.

botaniques. We did not use all of the nearly five million observations contained in the SILENE database at the date of the export (June 2016). As our analysis is based on a 5×5 km² grid cells, we decided to only retain the data whose georeferencement is precise enough. Furthermore, the distribution of several taxa is still insufficiently known and could distort the results. The final dataset results in 4,263,734 vegetation plant samples corresponding to 3,697 plant species. We divided the study area using a UTM grid composed of 2,607 squares of lateral size $l = 5$ km. In order to assess the impact of the spatial resolution on the results (Divíšek *et al.*, 2016; Lennon *et al.*, 2001), we also applied the aforementioned biogeographical network analysis with a grid composed of squares of lateral size $l = 10$ km (see Figure S1 and Table S1 in Appendix for more details).

Biogeographical network analysis

1. Biogeographical bipartite network. Delineating biogeographical regions based on species distribution requires a link between the species studied and their spatial environment. This link is usually identified with presence-absence matrices where each row represents a geographical location and each column a species. According to the type and quality of data, but also to the research question, the species *relevé* can be aggregated both spatially or by group of

species. The region of interest is usually divided into grid cells, the resolution of which depends mostly on the size of the study area, the taxonomic group under study and the accuracy of the data. Another way of formalizing complex interactions between species and geographical locations is to build a biogeographical bipartite network. This bipartite network enables us to model relations between two disjoint sets of nodes, geographical locations and species (in our case), which are linked by the presence of a species (or a group of species) in a given location during a certain time window (Step 1 in Figure 2). This way of understanding complex interactions makes it possible to visualize and analyze complex spatio-ecological systems as a whole from individual interactions to local and global biogeographical properties.

2. Delineating biogeographical regions. The purpose here is to identify biogeographical regions, hence we project our biogeographical bipartite network on a spatial template (Step 2 in Figure 2). To do so we defined a metric to measure the similarity of species composition between geographical locations. Several measures based on beta diversity have been proposed to quantify the degree of (dis)similarity between geographical locations, typically taking into account the number of shared species between geographical locations (Koleff *et al.*, 2003; Wilson & Shmida, 1984). These measures are mostly based on presence-absence data and aim at quantifying species turnover

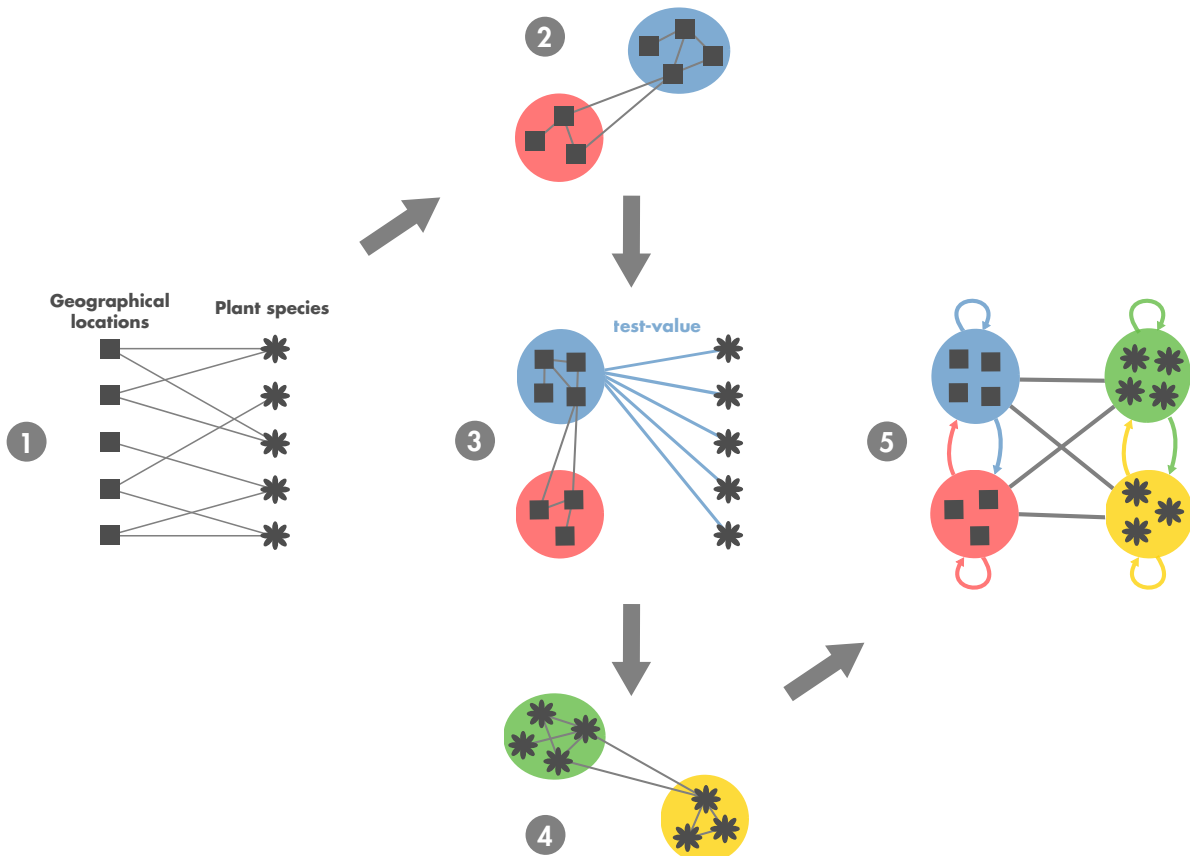


Figure 2. Steps of the biogeographical network analysis. 1. Biogeographical bipartite network where geographical locations and species are linked by the presence of a species (or a group of species) in a given location during a certain time window. Note that there is no link between nodes belonging to the same set. 2. The bipartite network is then spatially projected by using a similarity measure of species composition between geographical locations. Biogeographical regions are then identified with a network community detection algorithm. 3. The test-value matrix based on the contribution of species to biogeographical regions is computed. 4. Then, a network of similarity between species is built, based on the test-value matrix. Clusters of species sharing similar spatial features are identified using a community detection algorithm. 5. Finally, a coarse-grained biogeographical network unveiling the biogeographical structure of the studied area and the relationship between biogeographical regions is obtained.

and species nestedness among locations, together or separately (Baselga, 2012). Although this indicator may be influenced by gradients in species richness (Baselga, 2012; Dapporto *et al.*, 2015; Lennon *et al.*, 2001), better results were obtained with the Jaccard index in our case.

The resulting network is a weighted undirected spatial network whose intensity of links between geographical locations are range from 0, absence of a link because there are no species in common, to 1, i.e. identical species composition. The detection of community structure in biogeographical networks is an interesting alternative approach to delineating biogeographical regions (Kougioumoutzis *et al.*, 2014; Vilhena & Antonelli, 2015). Community structure is indeed an important feature, revealing both the network internal organization and similarity patterns among its individual elements. In this study we used the OSLOM algorithm (Lancichinetti *et al.*, 2011). OSLOM detects statistically significant

communities with respect to a global null model (i.e. random graph without community structure). The community statistical significance is measured as the probability of finding the community in a random null model. This algorithm is nonparametric in the sense that it identifies the statistically significant partition, without defining the number of communities *a priori*.

3. Test-value matrix. To analyse the biogeographical regions and their species composition, we rely on *test-values* measuring the under- or over-representation of species in a bioregion. Let us consider a studied area divided into n geographical locations, a species i present in n_i locations and a biogeographical region j composed of n_j geographical locations. The test-value compares the actual number of geographical locations n_{ij} , located in biogeographical region j and supporting species i , with the average number $n_i n_j / n$ that would be expected if the species were uniformly distributed over the whole studied

area. Since this quantity depends on n_i and n_j it is normalized by the standard deviation associated with the average expected number of geographical locations (Lebart *et al.*, 2000). The test value ρ_{ij} is then defined as,

$$\rho_{ij} = \frac{n_{ij} - \frac{n_i n_j}{n}}{\sqrt{\frac{n-n_j}{n-1} \left(1 - \frac{n_j}{n}\right) \frac{n_i n_j}{n}}} \quad (1)$$

The test value ρ_{ij} is negative if the species i is under-represented in region j , equal to 0 if the species i is present in region j in the same proportion as in the whole study area or positive if the species i is over-represented in region j . In the latter case we consider that the species i contribute positively to region j and the level of contribution depends of the ρ_{ij} value. Additionally, we consider that a plant species contribute positively and significantly to a bioregion j if ρ_{ij} is higher than a predetermined significance threshold α . Hence, The test-value matrix ρ can be used to highlight set of species which better characterize the biogeographical regions. The test-values are easy to interpret by specialists and represent an user-friendly way of ranking species according to their relevance.

4. Clusters of species. The next step is to identify how similarities between species are spatially distributed across the study area. Here also we build a network in which the similarity $s_{ii'}$ between two species i and i' is equal to,

$$s_{ii'} = \frac{1}{1 + \sqrt{\sum_j (\rho_{ij} - \rho_{i'j})^2}} \quad (2)$$

This similarity metric is based on the Euclidean distance between test-values for each pair of species. Again, the community detection algorithm OSLOM is used to detect significant groups of species sharing the same spatial features (Step 4 in Figure 2). This step produces a preliminary delimitation of the relationships between biogeographical regions by identifying how the clusters of species contributes to one or several biogeographical regions.

5. Coarse-grained biogeographical network. Several useful metrics can be derived from ρ to characterize relationships between biogeographical regions in a more quantitative manner. First we retained only the positive and significant species contributions by considering only test-values higher than $\alpha = 1.96$ (2.5% significance level of a Gaussian distribution).

$$\rho_{ij}^+ = \rho_{ij} \mathbb{1}_{\rho_{ij} > 1.96} \quad (3)$$

Then, since we are interested in interactions between biogeographical regions we focused on the way species contributions are distributed among regions by normalizing ρ^+ by row (Equation 4).

$$\hat{\rho}_{ij}^+ = \rho_{ij}^+ / \sum_i \rho_{ij}^+ \quad (4)$$

We then determined for each biogeographical regions j how the set of species $A_j = \{i \mid \rho_{ij} > 1.96\}$ that contributes to this biogeographical region are specific to it or also contribute to other regions (Equation 5).

$$\lambda_{jj'} = \frac{1}{|A_j|} \sum_{i \in A_j} \hat{\rho}_{ij'}^+ \quad (5)$$

The *specificity* of a biogeographical region is therefore measured with λ_{jj} , while the *relationships* with other regions is given by $\lambda_{jj'}$. It is important to note that for a given region j the vector λ_j . sum to one and can be expressed in percentage.

At the end of the process, we obtain a coarse-grained biogeographical network summarizing the biogeographical structure of the study area. This network is composed of the biogeographical regions and the species clusters (Step 5 in Figure 2). All the metrics used to measure the similarity between the different biogeographical regions are derived from the matrix of test-value ρ .

RESULTS

Biogeographical bipartite network

The bipartite network extracted from the database is composed of 2,607 5×5 km² grid cells and 3,697 plant species, where the links represent the occurrence of plant species in the grid cells. Two network degree distributions can be naturally associated to this type of network: the number of species per grid cell and the number of cells covered by each species. The probability density functions of these two distributions are displayed in Figure 3.

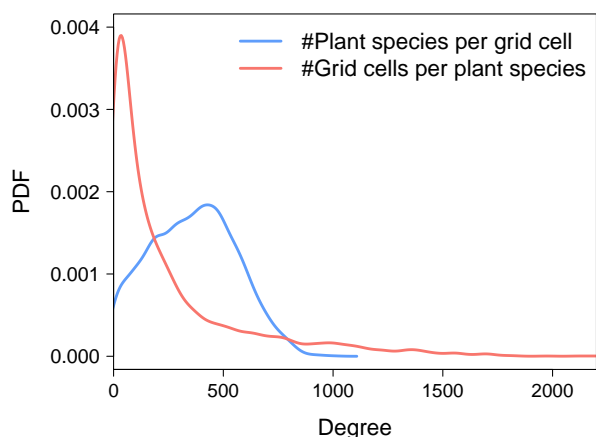


Figure 3. Degree distributions of the biogeographical bipartite network. Probability density function of the number of plant species per grid cell (in blue) and the number of cells covered per plant species (in red).

The spatial component of the network is very dense. Most of the grid cells host between 200 and 500 plant

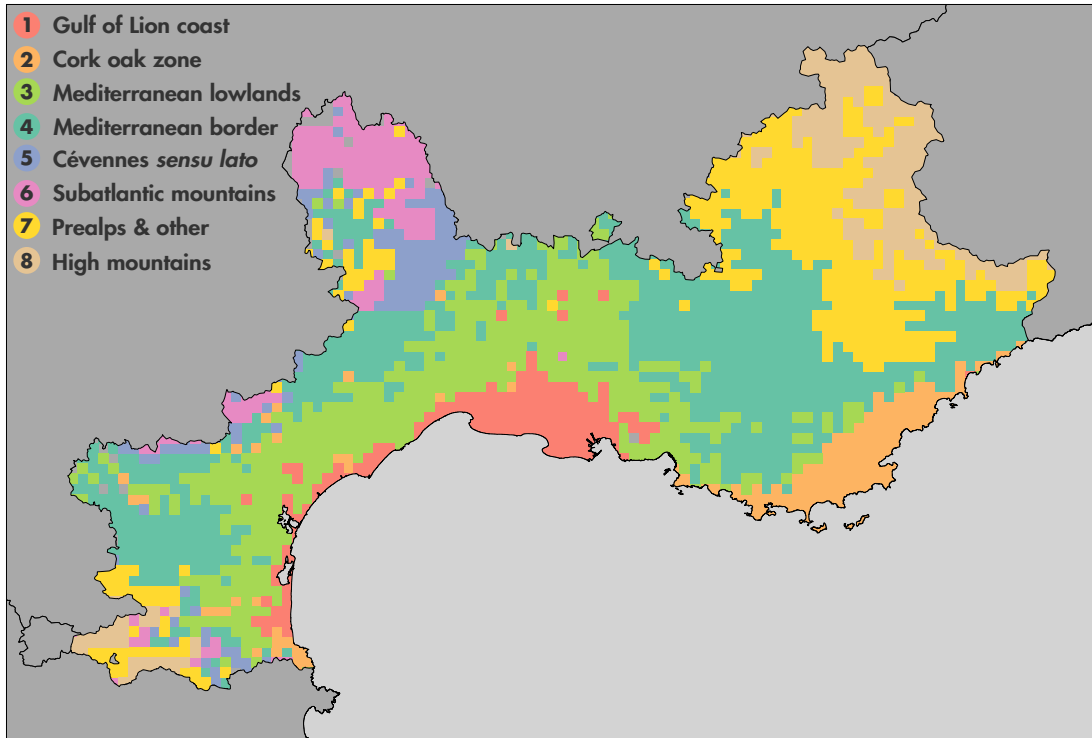


Figure 4. Biogeographical regions based on similarity in plant species ($l=5$ km). Eight bioregions have been identified. 1. Gulf of Lion coast in red. 2. Cork oak zone in orange. 3. Mediterranean lowlands in light green. 4. Mediterranean border in dark green. 5. Cévennes *sensu lato* in purple. 6. Subatlantic mountains in pink. 7. Prealps and other medium mountains in yellow. 8. High mountains in brown.

species, with an average of 360 species per cell (i.e. 15 species/km²). For species side, the situation is different; the majority of plant species cover less than 10% of the study area. However, the distribution exhibits a long tail with a non-negligible number of widespread species.

Delineating biogeographical regions

We identified eight statistically significant biogeographical regions reflecting the biogeographical structure of the French Mediterranean area based on plant species distribution (Figure 4). Clusters size vary from 120 to 807 square cells of lateral size $l = 5$ km. The clusters are spatially coherent, exhibiting a connectivity measure, ratio between the number of grid cells in the largest patch and the total number of grid cells (Turner *et al.*, 2001), always higher than 0.5. Note that the results obtained are not scale sensitive. The spatial coherence of each cluster according to the scale ($l = 5$ and 10 km) can be found in Table S1 in Appendix.

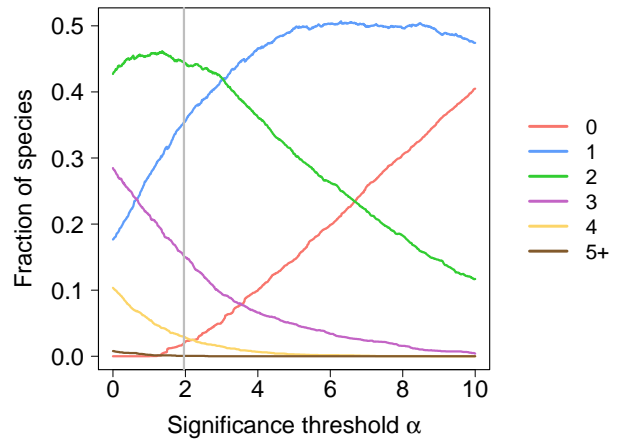


Figure 5. Fraction of species contributing positively and significantly to a given number of biogeographical regions (from 0 to 5 or more) as a function of the significance threshold. The vertical line represents the significance threshold $\alpha = 1.96$.

Clusters of plant species

The test-value matrix can be used to identify plant species that contribute positively to the different biogeographical regions. It is worth noting that the num-

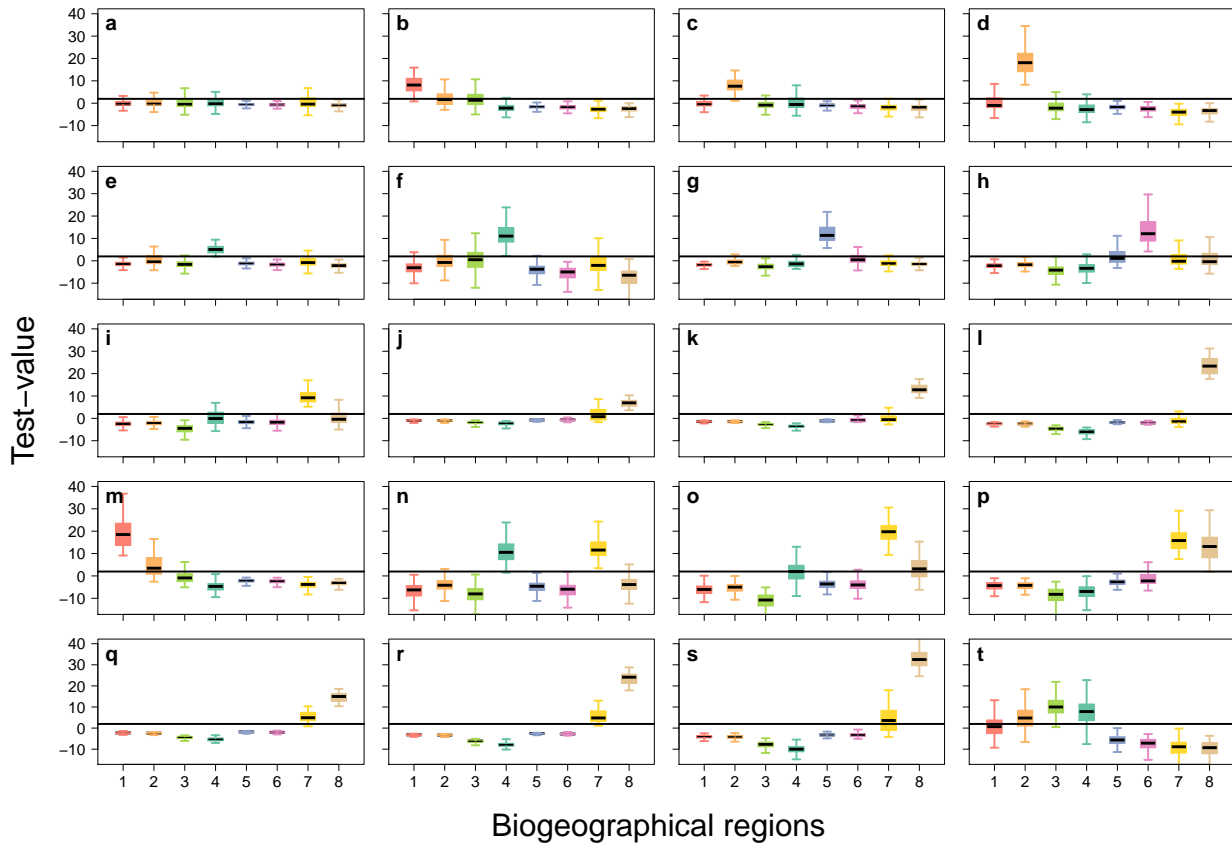


Figure 6. Description of the clusters of plant species. Boxplot of test-values according to the bioregions and the plant species clusters. The horizontal line represents the significance threshold $\alpha = 1.96$. The number of plant species per cluster is available in Table S2 in Appendix.

ber of contributions and their intensity vary among species. Indeed, some species contribute very little to only one region while other species contribute significantly to three or more regions. Figure 5 displays the fraction of species contributing positively to a given number of biogeographical regions as a function of a significance threshold value. If we consider a threshold $\alpha = 1.96$, that corresponds to a 2.5% significance level of a Gaussian distribution, we observe that the vast majority of plant species contributes positively to one or two regions representing 35% and 45% of species, respectively. There is also 20% percent of plant species that contribute to three or more biogeographical regions. If we increase the minimum level of contribution necessary to claim that a species contributes to a region, we see that the fraction of species contributing to two or more bioregions dramatically decreases while the fraction of species with no contribution increases. However, it is interesting to note that the fraction of species contributing to one region to increases until reaching a plateau. This demonstrates that 50% of plant species are strongly connected to a single region.

The similarities between plant species' contribution to the 8 regions allowed us to identify 20 clusters of species. A description of the plant species clusters

based on their contribution to each biogeographical region is displayed in Figure 6. We observed different patterns of contributions in terms of shape and intensity. This allows for the identification of clusters of species sharing similar spatial features and highlights relationships between biogeographical regions through the way plant species contribute to different group of regions.

Relationships between biogeographical regions

This leads us to the study of relationships between biogeographical regions. The network of interactions λ derived from the test-value matrix is plotted in Figure 7. We found that, globally, plant species contributing significantly to a region contribute mostly to this region, with an average specificity of 51% across the eight biogeographical regions. It must be pointed out however that some regions are more specific than others with λ_{jj} values ranging from 40% to 65%.

Analysis of how biogeographical regions connect with each other showed that there is no isolated region in the sense that every region is connected with at least one other region with a $\lambda_{jj'}$ value varying from 1 to 28%. Moreover, for all regions, the maxi-

mal $\lambda_{jj'}$ value is always higher than 10%. Although it is generally the case, it is also worth mentioning that the relationships are not necessarily symmetric. It represents an interesting way of detecting hierarchical relationships. A table displaying all $\lambda_{jj'}$ values is available in Table S3 in Appendix.

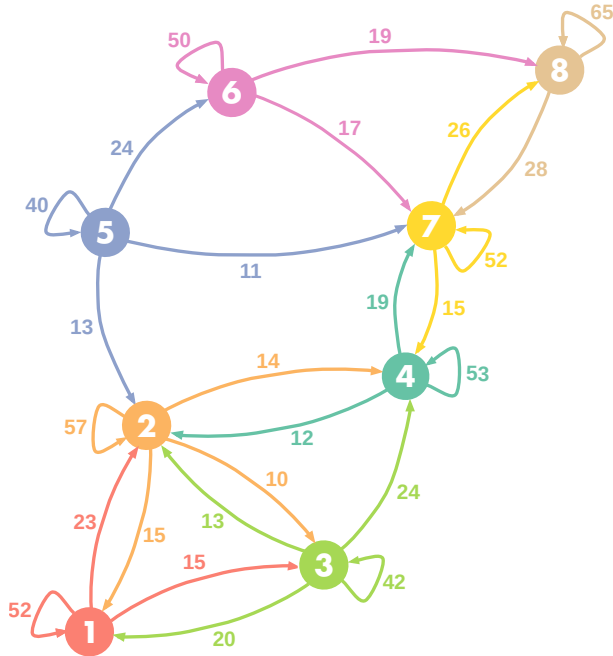


Figure 7. Network of interactions between biogeographical regions. Values are expressed in percentage. Only links with a value $\lambda_{jj'}$ higher than 10% are shown.

DISCUSSION

In this study we delineate spatial bioregions in southern France, a transition area between a Mediterranean and temperate climate. The present analysis represents to our knowledge one of the largest network-based studies published to date; we classified more than four million data points across a territory of about 558,776 km². We delineated eight statistically significant biogeographical regions, which we will first present in relation to previously published work, and emphasize their specificity regarding associated clusters of species. We discuss the observed spatial patterns in terms of ecological and historical drivers, to provide insights into mechanisms driving the assemblage of vegetation communities.

Biogeographical regions

The clustering approach identified eight statistically significant spatial clusters, that represents coherent territories detailed below. Regions are presented from Mediterranean toward temperate and

mountainous climates.

1. Gulf of Lion coast is a bioregion that extends west of the Rhône, penetrating more inland around the wetland of the Rhône Delta. The latter, along with the Languedoc lagoons, is frequently used as an example of azonal vegetation (Ozenda, 1994), and the originality of the flora and the vegetation of these areas has long been recognized (Molinier & Tallon, 1970). Some subdivisions have been suggested separating, even at a coarse scale, the sand-dune complex, the halophytic vegetation and the salt meadows (Bohn *et al.*, 2000), but were not found here probably due to the scale at which we worked and thus the precision of the cells we used. From a geological point of view, this bioregion is essentially made of sand dunes, lagoon sediments and modern alluvium. It is entirely situated under a Mediterranean climate, in the mesomediterranean climatic belt, with a dry season of two or three months in the summer (Rivas-Martínez *et al.*, 2004a). It is important to note that taxa specific to this cluster exhibit a distribution following the Mediterranean coastal area, extending in some cases towards other coastal areas or to arid inland zones. They are mostly encountered in halophytic communities and surprisingly not that much in dune ones, suggesting that the key factor defining this bioregion might be the saline soils rather than the coastal position alone.

2. Cork oak zone encompasses the Maures-Estérel range and neighbouring areas. West of the Rhône, it is fragmented with cells in the eastern tip of the Pyrenees (low Albères and the Roussillon lowlands), plus a few more sparsely dispersed zones in Languedoc. The Provence and Albères areas have been identified by phytogeographers (Ozenda, 1994; Ozenda & Lucas, 1987) as the cork oak zone, a silicicolous warm mesomediterranean area. Indeed, climatic data show a clear summer dry period of one to two months. Almost all of the cells contain acidic soils over a variety of substrates (granites, gneiss, schists, sandstones, alluvial deposits, etc.). Species most linked to the “Cork oak zone” have a Mediterranean distribution, with some extending towards the Atlantic area. Characteristic species have ecological preferences for acid soils, and belong to various vegetation stages (forest, scrub or grassland formations).

3. Mediterranean lowlands bioregion covers the hinterland of the Gulf of Lion from the Roussillon to western Provence. Several authors have individualized an arc shaped mesomediterranean zone (Dupias & Rey, 1985; Ozenda, 1994) but their limits do not fit exactly ours. The closest match is the catalonian provençal mesomediterranean holm oak forests unit of the European natural vegetation map (Bohn *et al.*, 2000). The area is principally composed of sedimentary rocks (mostly limestones and marls) and alluvium. Its climate is Mediter-

ranean, entirely situated in the mesomediterranean belt (Rivas-Martínez *et al.*, 2004a), with a summer dry period of one to three months. With few exceptions, species most linked to this bioregion have a distribution included in the Mediterranean region (Rivas-Martínez *et al.*, 2004b). A part of them belongs to communities of the *Quercetea ilicis* or of the former *Thero-Brachypodietae*, i.e. the matorral / forest and grasslands communities making up the landscape locally called “garrigues”. The other part of these taxa is usually found in disturbed communities, showing the strong incidence of human activities in this area.

4. Mediterranean border is a bioregion that follows the limit of the Mediterranean world as it is usually depicted (Dupias & Rey, 1985; Quézel & Médail, 2004) and corresponds roughly to what has been called a supramediterranean belt (Ozenda, 1994) or a submediterranean zone (Bólos, 1961). It fits quite well with four mapping units of the Map of the natural vegetation of Europe (Bohn *et al.*, 2000); namely the catalonian provençal supramediterranean holm oak forests and three types of downy oak forests (ligurian middle apennine, languedocian and those extending from the southern Pyrenees to the southwest pre-Alps). The substratum of this area is mainly calcareous and marly. This area has a short (one month) summer drought period with the exception of some Var and Alpes-Maritimes places where the summer drought is more pronounced (two months). The global distribution of species most linked to this bioregion is centred on the western Mediterranean, sometimes going further north or east, which corresponds to a western eury-mediterranean distribution. A large part of species share a common ecology, occurring frequently in communities belonging to the *Helianthemo italicici* - *Aphyllanthion monspeliensis* and to a lesser extent to the *Ononidetalia striatae* (Gaultier, 1989; Rivas-Martínez *et al.*, 2002), i.e. dry dwarf scrubs and their associated grasslands on calcareous and marly eroded soils (Mucina *et al.*, 2016).

5. Cévennes sensu lato is a bioregion to which most of the cells are situated in the Cévennes areas, while the remainder is scattered over the eastern Pyrenees piedmont and the Montagne Noire (southern limit of the Massif Central). This spatial cluster overlays four zones of the phyto-ecological regions (Dupias & Rey, 1985): the lower Cévennes, the “warm” Cévennes valleys, the Aspres and the chestnut zone of the southern edge of the Montagne Noire. The Cévennes proper part of this cluster has also been identified by other authors (Braun-Blanquet, 1923; Ozenda, 1994) and putative glacial refugia has been positioned there (Médail & Diadema, 2009). This area is not subject to a summer drought and covers siliceous substrata such as schists, granites or gneiss. Taxa exhibiting the strongest link to this biogeographical region are either Cévennes endemics,

subendemics (Dupont, 2015; Lavergne *et al.*, 2004) or plants with a more or less Atlantic distribution (Dupont, 2015). Yet, there is no clear emerging ecological pattern among these taxa.

6. Subatlantic mountains The largest area covered by cells of this biogeographical region is the northern part of the Lozère department. The remaining cells are mostly distributed in the Massif Central in the following ranges: Mont Lozère, Mont Aigoual, Haut-Languedoc, and Montagne Noire. A few more patches can be found in the Pyrenees: in the Capcir, the High Vallespir and the Albères ridge. These areas belong to the beech montane belt (Bohn *et al.*, 2000; Ozenda, 1994) with a few exceptions where Scots pines dominate. It corresponds to the predominantly siliceous subatlantic type (Ozenda & Lucas, 1987), where the climate is rather wet, with precipitations frequently exceeding 1,000 mm per year and no dry period. Thus, wetlands and bogs are not rare, and the substratum is made of igneous rocks which explain the acidic nature of the soils. The majority of the taxa most linked to this spatial cluster are generally distributed all over the eurosiberian region or the western part of this region, corresponding to a subatlantic distribution (Dupont, 2015; Rivas-Martínez *et al.*, 2004b). An interesting point is that the majority of those plants are growing in wetlands habitats, a trend already noticed in the Massif Central (Braun-Blanquet, 1923).

7. Pre-Alps and other medium mountains represent a bioregion whose cells are disseminated through the lower parts of the eastern Pyrenees including almost all the Pyrenean part of the Aude department, through the highest areas of the Causses, around the Mont Ventoux and through the most eastern part of the Pre-Alps. This area has rarely been individualised in such a way even if at a European scale it can be related to several more or less calcicolous beech or fir-beech forest belts (Bohn *et al.*, 2000), or more specifically, for the Var department, to a pre-alpine district (Lavagne, 2008). Most of the rock underlying this area is calcareous. Climatically, we are outside of the Mediterranean climate as there is no dry period. The distribution of taxa most linked to this biogeographical region is basically holarctic, avoiding the Mediterranean parts of Europe. Some of these taxa also avoid the most Atlantic part of the continent. Their ecology is varied, pertaining to different stages (grasslands, shrubs, forests) of mountain vegetation series, often (but not systematically) calcicolous.

8. High mountains This bioregion consists of the highest part of the Alps and the Pyrenees in the study region. If most authors agree on individualizing the upper vegetation belts of these mountain ranges, its unity and the common points are less often identified (Ozenda, 2002). Both calcareous and acidic soils are to be found in this area. Cells of this region are the

coldest of our study area, and there is no dry period: the climate is relatively harsh and the vegetation period is reduced (Ozenda, 2002) compared to what happens in the other clusters. Taxa most linked to this region are mainly European mountains endemics, venturing also in the Arctic. They all belong to grasslands or snowbeds communities, which is consistent with their occurrence on the highest ranges.

Species and spatial relationships among biogeographical regions

Defining the Mediterranean region

At a global scale, the delimitation of the Mediterranean border has been a long running question (Latini *et al.*, 2017), and the imprecision and mismatch of the numerous attempts attest to the difficulties. The first attempt goes back to third edition of the Flore Française by Lamarck & Candolle (1805) in Ebach & Goujet (2006) followed by several other works such as Ozenda (1994), or Rivas-Martínez *et al.* (2004a). However, variability in results do not lead to a comprehensive framework. This has several implications regarding conservation programs, as the delimitation mentioned by European legislation has been used as a reference to delimit the distribution of several protected habitat². In this study, the network approach allowed to discriminate two “sub-networks” with little exchange regarding species composition and different relative contribution to each area. They globally relate to a temperate and a Mediterranean sub-groups, whose main climatic difference relies on the existence of a summer dry period. The absence of orogenic barriers along this climate-based distinction is likely to produce shallow boundaries typical of transition areas (Antonelli, 2017; Ficetola *et al.*, 2017) exemplified here by the cluster “Mediterranean border”. Thus, instead of drawing a clear line, we propose to identify a transition area (Droissart *et al.*, 2017; Latini *et al.*, 2017) with an upper boundary as the limit of the Mediterranean biome (Antonelli, 2017).

Vicariance and fragmentation among bioregions

At a finer scale, the relationship between biogeographical regions can be seen through the understanding of species relative importance in each area. First, the regions “Gulf of Lion coast”, “Cork oak zone” and “Mediterranean lowlands” are all included within the same bioclimatic belt (Rivas-Martínez *et al.*, 2004a). Their differentiation mostly relies on substratum, i.e.

calcareous (bioregion 3), siliceous (bioregion 2) or quaternary deposits (bioregion 1). Thus, they are well defined and little uncertainty exists concerning their spatial configuration (Figure S2 in Appendix); those three entities can be seen as vicariant climatic entities which have conjointly developed on different geological substrates, or “islands”. As a result, they share an important pool of species, and present the highest complementarity in the network, as they are the only three clusters all related to each other. In contrast, the relationship between the “Cork oak zone” and the “Cévennes” exemplify the opposite process: those two areas share a similar bedrock (mainly acidic substrate) but are located at each extreme of the Mediterranean climatic gradient. While the “Cork oak zone” is present under hot and dry mesomediterranean climate (some coastal cells even belonging to the thermomediterranean belt), the “Cévennes” present a higher pluviometry and a very weak summer drought. Consequently, they share a common set of species which interestingly are typical of the “Cévennes” cluster, and extend into the “Cork oak zone”. Noteworthy point, those population can constitute relictual rear edge populations, which often retain particular interest for conservation (Hampe & Petit, 2005; Lavergne *et al.*, 2006).

Finally, the “Pre-Alps” and “High mountains” bioregions are both present within the three mountain chains, and occupy climatic conditions with no dry period at all, and especially harsh prolonged winter for the second. Several species clusters are highly informative for both of those bioregions, which signify that they share an important group of species globally adapted to mountain environment. “High mountains” present the highest percentage of typical species. Yet, within the numerous plant species clusters characterizing those entities (5 clusters in Figure 6), the relative contribution of each toward one or the other bioregion might differ slightly, sometimes in association with another bioregion such as the “Mediterranean border” (Figure 6). This illustrates that groups of taxa are unevenly important across these two regions, probably reflecting the complex geological substrate. Thus, while our analysis reflect an overall homogeneity of mountain flora mainly driven by climate, it is likely that finer divisions based on a more precise study could be expected.

Eco-evolutionary factors driving the spatial organisation of plant diversity

The spatial distribution and species relative importance for each region are related to the processes that have shaped Mediterranean biota in the south of France. Among ecological factors playing a role in setting those limits, orographic barriers and past tectonic movement are expected to have little impact on our study area, as no such events have occurred since the onset of the Mediterranean climate in the Pliocene (Rosenbaum *et al.*, 2002). Therefore, spatial

² http://ec.europa.eu/environment/nature/info/pubs/docs/biogeos/Mediterranean/KH7809610FRC_002.pdf

structuration relies on two elements. On the one hand, a climatic gradient from Mediterranean to temperate creates fuzzy spatial limits among adjacent groups, and increases uncertainty when delimitating groups (Figure S2 in Appendix). This is exemplified by the spatial imbrication of “Mediterranean lowlands” and “Mediterranean border”. On the other hand, geological variations can form sharp transitions creating important species turnover between places close apart. This is exemplified by the “Cork oak zone” whose spatial delimitation is very clear, due to the presence of an acidic substrate while surrounding places are dominated by calcareous-based rock. Interestingly, this area still shares an important part of its biota with other places in the Mediterranean basin probably inherited from times where such geological islands formed a single ensemble, before the separation and later migration of these islands (Médail & Quézel, 1997; Rosenbaum *et al.*, 2002).

This joint action of these two ecological factors has already been highlighted in previous bioregionalization of the Mediterranean basin (Buirra *et al.*, 2017). As a result, understanding spatial patterns provides two insights into the history of Mediterranean plant diversity. First, complex geo-climatic variation have played a key role in shaping island-like territories which have fragmented species distributions, a factor that has strong influence on populations characteristics both genetically and demographically (Pironon *et al.*, 2017). This is a key characteristic of the Mediterranean basin flora which shows recurrent patterns of narrow endemism, species turnover and highly disjunct distributions (Thompson, 2005).

However, this does not identify the reason why a given taxon is present, could it be recent dispersal following glacial cycle (Lumaret *et al.*, 2002), or long term persistence in a given refugia (Médail & Didema, 2009; Papuga *et al.*, 2015). Second, the association between spatial isolation and ecological originality is key to understand speciation patterns in the Mediterranean. While allopatric isolation has been suspected to be the main mechanism explaining the differentiation of taxa, the shared significance of different ecological variables (namely climate and geology) points out the combined importance of spatial isolation and heterogeneous selection pressures (Anacker & Strauss, 2014; Thompson, 2005). Additionally, recent studies have shown that this can be enhanced by small scale changes of the ecological niche (Papuga *et al.*, 2018), and thus favour the emergence of narrow range endemic species, a key feature of the Mediterranean flora (Lavergne *et al.*, 2004; Thompson *et al.*, 2005). Altogether, this bioregionalization set the scene to understand the shared phylogeographic legacy of the Mediterranean biota (Puşças & Choler, 2012).

Network approaches in biogeography

The quality of a bioregionalization is dependent on the data and the method used. To our knowledge,

the present analysis constitutes the densest species-cells network analysed in a bioregionalization study, at such a high spatial resolution. Therefore, results of this study demonstrate that new statistical methods based on network analysis can bring solutions to manage and analyze large databases, and provide efficient bioregionalization. Additionally, they present the important advantage of allowing parallel analysis of both taxa and bioregions, a key step in our understanding of those patterns (Morrone, 2018). However, the application of network methods is not an absolute perfect method, and some methodological gaps remain open. We attempted to use the map equation (Rosvall & Bergstrom, 2008) directly on the bipartite network as in Vilhena & Antonelli (2015), but it failed to detect any coherent biogeographical regions. This is probably due to the high density of plant species records within each cell. However, it is worth mentioning that despite the high quality of available dataset, the study area has not been evenly inventoried, which supposes that an increase in the number of records per species is to be expected in the future. While this opens new paths for bioregionalization (including downscaling cell size to improve precision of maps), statistical methods remain inexistent to treat those questions.

ACKNOWLEDGEMENTS

This work was supported by a grant from the French National Research Agency (project NetCost, ANR-17-CE03-0003 grant). Partial financial support has been received from the French Ministère de la Transition Ecologique et Solidaire (MTES). We thank the Alpine National Botanic Conservatory for providing some of the Provence-Alpes-Côte d’Azur data. We wish to thank Christelle Hély-Alleaume and Virgile Noble for useful discussions. A special thank goes to John D. Thompson for correcting English and interesting remarks.

DATA AVAILABILITY

The test-value matrix is available at <https://github.com/maximelenormand/Biogeographical-network-analysis>.

REFERENCES

- Anacker, B. L. & Strauss, S. Y. (2014) The geography and ecology of plant speciation: range overlap and niche divergence in sister species. *Proceedings of the Royal Society of London B: Biological Sciences*, **281**, 20132980.
- Antonelli, A. (2017) Biogeography: Drivers of bioregionalization. *Nature Ecology & Evolution*, **1**, 0114.

- Baselga, A. (2012) The relationship between species replacement, dissimilarity derived from nestedness, and nestedness. *Global Ecology and Biogeography*, **21**, 1223–1232.
- Blondel, J., Aronson, J., Bodiou, J.-Y. & Boeuf, G. (2010) *The Mediterranean Region: Biological Diversity in Space and Time*. Oxford University Press, Oxford, New York, second edition edition.
- Bohn, U., Gollub, G. & Hettwer, C. (2000) *Karte der natürlichen Vegetation Europas*. Bundesamt für Naturschutz. Landwirtschaftsvlg Münster, Bonn, 1., Aufl. edition.
- Bólos, O. (1961) La transición entre la depresión del ebro y los pirineos en el aspecto geobotánico. *Anal. Inst. Bot. Cavanilles*, **18**, 99–254.
- Braun-Blanquet, J. (1923) *L'origine et le développement des flores dans le massif central de France; avec aperçu sur les migrations des flores dans l'Europe sud-occidentale*. L. Lhomme, Paris.
- Buira, A., Aedo, C. & Medina, L. (2017) Spatial patterns of the Iberian and Balearic endemic vascular flora. *Biodiversity and Conservation*, **26**, 479–508.
- Cheruvilil, K. S., Yuan, S., Webster, K. E., Tan, P.-N., Lapierre, J.-F., Collins, S. M., Fergus, C. E., Scott, C. E., Henry, E. N., Soranno, P. A., Filstrup, C. T. & Wagner, T. (2017) Creating multithemed ecological regions for macroscale ecology: Testing a flexible, repeatable, and accessible clustering method. *Ecology and Evolution*, **7**, 3046–3058.
- Dahlin, K. M., Asner, G. P. & Field, C. B. (2014) Linking vegetation patterns to environmental gradients and human impacts in a mediterranean-type island ecosystem. *Landscape Ecology*, **29**, 1571–1585.
- Dapporto, L., Ciolli, G., Dennis, R. L. H., Fox, R. & Shreeve, T. G. (2015) A new procedure for extrapolating turnover regionalization at mid-small spatial scales, tested on British butterflies. *Methods in Ecology and Evolution*, **6**, 1287–1297.
- Divíšek, J., Storch, D., Zelený, D. & Culek, M. (2016) Towards the spatial coherence of biogeographical regionalizations at subcontinental and landscape scales. *Journal of Biogeography*, **43**, 2489–2501.
- Droissart, V., Dauby, G., Hardy, O. J., Deblauwe, V., Harris, D. J., Janssens, S., Mackinder, B., Blach-Overgaard, A., Sonké, B., Sosef, M. M., Stévant, T., Svenning, J.-C., Wieringa, J. J. & Couvreur, T. L. P. (2017) Beyond trees: Biogeographical regionalization of tropical Africa. *Journal of Biogeography*.
- Dupias, G. & Rey, P. (1985) *Document pour un zonage des régions phyto-écologiques*. Centre d'écologie des ressources renouvelables.
- Dupont, P. (2015) *Les plantes vasculaires atlantiques, les pyrénéo-cantabriques et les éléments floristiques voisins dans la péninsule ibérique et en France*. Société botanique du Centre-Ouest.
- Ebach, M. C. & Goujet, D. F. (2006) The first biogeographical map. *Journal of Biogeography*, **33**, 761–769.
- Fenu, G., Fois, M., Cañadas, E. M. & Bacchetta, G. (2014) Using endemic-plant distribution, geology and geomorphology in biogeography: the case of Sardinia (Mediterranean Basin). *Systematics and Biodiversity*, **12**, 181–193.
- Ficetola, G. F., Mazel, F. & Thuiller, W. (2017) Global determinants of zoogeographical boundaries. *Nature Ecology & Evolution*, **1**, 0089.
- Funk, V. A., Richardson, K. S. & Sakai, A. K. (2002) Systematic Data in Biodiversity Studies: Use It or Lose It. *Systematic Biology*, **51**, 303–316.
- Gaultier, C. (1989) *Relations entre pelouses eurosibériennes (Festuco-Brometea Br. -Bl. Et Tx. 43) et groupements méditerranéens (Ononido-Rosmarinetea Br. -Bl. 47) : étude régionale (Diois) et synthèse sur le pourtour méditerranéen Nord-occidental*. PhD thesis, Université de Paris-Sud.
- Graham, C. H. & Hijmans, R. J. (2006) A comparison of methods for mapping species ranges and species richness. *Global Ecology and Biogeography*, **15**, 578–587.
- Greuter, W. (1991) Botanical diversity, endemism, rarity and extinction in the Mediterranean area: an analysis based on the published volumes of MedChecklist. *Botanika Chronika*, **10**, 63–79.
- Guimerà, R. & Nunes Amaral, L. A. (2005) Functional cartography of complex metabolic networks. *Nature*, **433**, 895–900.
- Hampe, A. & Petit, R. J. (2005) Conserving biodiversity under climate change: the rear edge matters. *Ecology Letters*, **8**, 461–467.
- He, J., Kreft, H., Gao, E., Wang, Z. & Jiang, H. (2016) Patterns and drivers of zoogeographical regions of terrestrial vertebrates in China. *Journal of Biogeography*.
- Heikinheimo, H., Fortelius, M., Eronen, J. & Mannila, H. (2007) Biogeography of European land mammals shows environmentally distinct and spatially coherent clusters. *Journal of Biogeography*, **34**, 1053–1064.

- Joly, D., Brossard, T., Cardot, H., Cavailhes, J., Hilal, M. & Wavresky, P. (2010) Les types de climats en France, une construction spatiale. *Cybergeo : European Journal of Geography*.
- Koleff, P., Gaston, J. & Lennon, J. (2003) Measuring beta diversity for presence-absence data. *Journal of Animal Ecology*, **72**, 367–382.
- Kougioumoutzis, K., Simaiakis, S. M. & Tiniakou, A. (2014) Network biogeographical analysis of the central Aegean archipelago. *Journal of Biogeography*, **41**, 1848–1858.
- Kreft, H. & Jetz, W. (2010) A framework for delineating biogeographical regions based on species distributions. *Journal of Biogeography*, **37**, 2029–2053.
- Lamarck, J. d. M. d. & Candolle, A. (1805) *Flore française, ou descriptions succinctes de toutes les plantes qui croissent naturellement en France, disposées selon une nouvelle méthode d'analyse, et précédées par un exposé des principes élémentaires de la botanique*. Paris.
- Lancichinetti, A., Radicchi, F., Ramasco, J. J. & Fortunato, S. (2011) Finding Statistically Significant Communities in Networks. *PLOS ONE*, **6**, e18961.
- Latini, M., Bartolucci, F., Conti, F., Iberite, M., Nicoletta, G., Scoppola, A. & Abbate, G. (2017) Detecting Phytogeographic Units Based on Native Woody Flora: A Case Study in Central Peninsular Italy. *The Botanical Review*, **83**, 253–281.
- Lavagne, A. (2008) Synthèse phytogéographique du département du Var. In *Le Var et sa flore. Plantes rares ou protégées*. Naturalia Publications, Inflovar.
- Lavergne, S., Molina, J. & Debussche, M. (2006) Fingerprints of environmental change on the rare mediterranean flora: a 115-year study. *Global Change Biology*, **12**, 1466–1478.
- Lavergne, S., Thompson, J. D., Garnier, E. & Debussche, M. (2004) The biology and ecology of narrow endemic and widespread plants: a comparative study of trait variation in 20 congeneric pairs. *Oikos*, **107**, 505–518.
- Lebart, L., Piron, M. & Morineau, A. (2000) *Statistique exploratoire multidimensionnelle*. Dunod, Paris.
- Lennon, J. J., Koleff, P., Greenwood, J. J. D. & Gaston, K. J. (2001) The geographical structure of British bird distributions: Diversity, spatial turnover and scale. *Journal of Animal Ecology*, **70**, 966–979.
- Linder, H. P., de Klerk, H. M., Born, J., Burgess, N. D., Fjeldså, J. & Rahbek, C. (2012) The partitioning of Africa: Statistically defined biogeographical regions in sub-Saharan Africa. *Journal of Biogeography*, **39**, 1189–1205.
- Lumaret, R., Mir, C., Michaud, H. & Raynal, V. (2002) Phylogeographical variation of chloroplast DNA in holm oak (*Quercus ilex* L.). *Molecular Ecology*, **11**, 2327–2336.
- Mateo, R. G., Vanderpoorten, A., Muñoz, J., Laenen, B. & Désamoré, A. (2013) Modeling Species Distributions from Heterogeneous Data for the Biogeographic Regionalization of the European Bryophyte Flora. *PLOS ONE*, **8**, e55648.
- Médail, F. & Diadema, K. (2009) Glacial refugia influence plant diversity patterns in the Mediterranean Basin. *Journal of Biogeography*, **36**, 1333–1345.
- Médail, F. & Quézel, P. (1997) Hot-Spots Analysis for Conservation of Plant Biodiversity in the Mediterranean Basin. *Annals of the Missouri Botanical Garden*, **84**, 112–127.
- Mikolajczak, A., Marchal, D., Sanz, T., Isenmann, M., Thierion, V. & Luque, S. (2015) Modelling spatial distributions of alpine vegetation: A graph theory approach to delineate ecologically-consistent species assemblages. *Ecological Informatics*, **30**, 196–202.
- Molinier, R. & Tallon, G. (1970) *Prodrome des unités phytosociologiques observées en Camargue*. Molinier.
- Moreno Saiz, J. C., Donato, M., Katinas, L., Crisci, J. V. & Posadas, P. (2013) New insights into the biogeography of south-western Europe: Spatial patterns from vascular plants using cluster analysis and parsimony. *Journal of Biogeography*, **40**, 90–104.
- Morrone, J. J. (2018) The spectre of biogeographical regionalization. *Journal of Biogeography*, **105**, 1118–1123.
- Mucina, L., Bütmann, H., Dierßen, K., Theurillat, J.-P., Raus, T., Čarni, A., Šumberová, K., Willner, W., Dengler, J., Garcí, R. G., Chytrý, M., Hájek, M., Di Pietro, R., Iakushenko, D., Pallas, J., Daniëls, F. J., Bergmeier, E., Santos Guerra, A., Ermakov, N., Valachovič, M., Schaminée, J. H. J., Lysenko, T., Didukh, Y. P., Pignatti, S., Rodwell, J. S., Capelo, J., Weber, H. E., Solomeshch, A., Dimopoulos, P., Aguiar, C., Hennekens, S. M. & Tichý, L. (2016) Vegetation of Europe: hierarchical floristic classification system of vascular plant, bryophyte, lichen, and algal communities. *Applied Vegetation Science*, **19**, 3–264.

- Murray, A. (1866) *The geographical distribution of mammals*. Day and Son, limited,, London,.
- Myers, N., Mittermeier, R. A., Mittermeier, C. G., da Fonseca, G. A. B. & Kent, J. (2000) Biodiversity hotspots for conservation priorities. *Nature*, **403**, 853–858.
- Ozenda, P. (1994) *Végétation du continent Européen*. Delachaux et Niestlé, Lausanne.
- Ozenda, P. (2002) *Perspectives pour une géobiologie des montagnes*. Presses Polytechniques et Universitaires Romandes, Lausanne.
- Ozenda, P. & Lucas, M. J. (1987) Esquisse d'une carte de la végétation potentielle de la France 1/1 500 000. *Documents de cartographie écologique*.
- Papuga, G., Gauthier, P., Pons, V., Farris, E. & Thompson, J. (2018) Ecological niche differentiation in peripheral populations: a comparative analysis of eleven Mediterranean plant species. *Ecography*, **176**, 724–738.
- Papuga, G., Gauthier, P., Ramos, J., Pons, V., Pironon, S., Farris, E. & Thompson, J. D. (2015) Range-Wide Variation in the Ecological Niche and Floral Polymorphism of the Western Mediterranean Geophyte *Narcissus dubius* Gouan. *International Journal of Plant Sciences*, **176**, 724–738.
- Pironon, S., Papuga, G., Villellas, J., Angert, A. L., García, M. B. & Thompson, J. D. (2017) Geographic variation in genetic and demographic performance: new insights from an old biogeographical paradigm. *Biological Reviews*, **92**, 1877–1909.
- Puşcaş, M. & Choler, P. (2012) A biogeographic delineation of the European Alpine System based on a cluster analysis of *Carex curvula*-dominated grasslands. *Flora - Morphology, Distribution, Functional Ecology of Plants*, **207**, 168–178.
- Quézel, P. (1999) Les grandes structures de végétation en région méditerranéenne: Facteurs déterminants dans leur mise en place post-glaciaire. *Geobios*, **32**, 19–32.
- Quézel, P. & Médail, F. (2004) *Ecologie et biogéographie des forêts du bassin méditerranéen*. Elsevier Masson, Paris.
- Ricklefs, R. E. (2004) A comprehensive framework for global patterns in biodiversity. *Ecology Letters*, **7**, 1–15.
- Rivas-Martínez, S., Díaz, T., Fernández-González, F., Izco, J., Loidi, J., Lousã, M. & Penas, A. (2002) Vascular plant communities of Spain and Portugal: addenda to the syntaxonomical checklist of 2001. Part II. *Itinera Geobot.*, **15**(2), 433–922.
- Rivas-Martínez, S., Penas, A. & Díaz, T. (2004a) *Bioclimatic Map of Europe*. University of León.
- Rivas-Martínez, S., Penas, A. & Díaz, T. (2004b) *Biogeographic Map of Europe*. University of León.
- Rosenbaum, G., Lister, G. S. & Duboz, C. (2002) Reconstruction of the tectonic evolution of the western Mediterranean since the Oligocene. *Journal of the Virtual Explorer*, **8**, 107–130.
- Rosvall, M. & Bergstrom, C. T. (2008) Maps of random walks on complex networks reveal community structure. *Proceedings of the National Academy of Sciences*, **105**, 1118–1123.
- Rueda, M., Rodríguez, M. & Hawkins, B. A. (2010) Towards a biogeographic regionalization of the European biota. *Journal of Biogeography*, **37**, 2067–2076.
- Rushton, S. P., Ormerod, S. J. & Kerby, G. (2004) New paradigms for modelling species distributions? *Journal of Applied Ecology*, **41**, 193–200.
- Stoddart, D. R. (1992) *Biogeography of the Tropical Pacific*. University of Hawaii Press.
- Tassin, C. (2017) *Paysages végétaux du domaine méditerranéen : Bassin méditerranéen, Californie, Chili central, Afrique du Sud, Australie méridionale*. Rfrence. IRD ditions, Marseille.
- Thompson, J. D. (2005) *Plant Evolution in the Mediterranean*. Oxford University Press, Oxford, New York.
- Thompson, J. D., Lavergne, S., Affre, L., Gaudeul, M. & Debussche, M. (2005) Ecological Differentiation of Mediterranean Endemic Plants. *Taxon*, **54**, 967–976.
- Tison, J.-M. & Foucault, B. d. (2014) *Flora Gallica : Flore de France*. Biotope Editions, Mze.
- Tison, J.-M., Jauzein, P., Michaud, H., Jeanmono, D. & Boillot, F. (2014) *Flore de la France méditerranéenne continentale*. Naturalia Publications, Turriers, 1re dition edition.
- Turner, M. G., Gardner, R. H. & O'Neill, R. V. (2001) *Landscape Ecology in Theory and Practice: Pattern and Process*. Springer-Verlag.
- Vilhena, D. A. & Antonelli, A. (2015) A network approach for identifying and delimiting biogeographical regions. *Nature Communications*, **6**, 6848.
- Wahlenberg, G. (1812) *Flora Lapponica*. Taberna libraria scholae realis, Berolini.
- Wallace, A. R. (1876) *The geographical distribution of animals : with a study of the relations of living and extinct faunas as elucidating the past changes of the earth's surface*. Harper and Brothers, Publishers, New York.

- Walter, H. & Breckle, S.-W. (1991) *Ökologie der Erde. Band 4: Spezielle Ökologie der Gemäßigten und Arktischen Zonen außerhalb Euro-Nordasiens*. Gustav Fischer Verlag.
- Walter, H. & Breckle, S.-W. (1994) *Ökologie der Erde. Band 3: Spezielle Ökologie der Gemäßigten und Arktischen Zonen außerhalb Euro-Nordasiens*. Gustav Fischer Verlag.
- Wilson, M. V. & Shmida, A. (1984) Measuring Beta Diversity with Presence-Absence Data. *Journal of Ecology*, **72**, 1055–1064.
- Zhang, M.-G., Slik, J. W. F. & Ma, K.-P. (2016) Using species distribution modeling to delineate the botanical richness patterns and phytogeographical regions of China. *Scientific Reports*, **6**, 22400.

APPENDIX

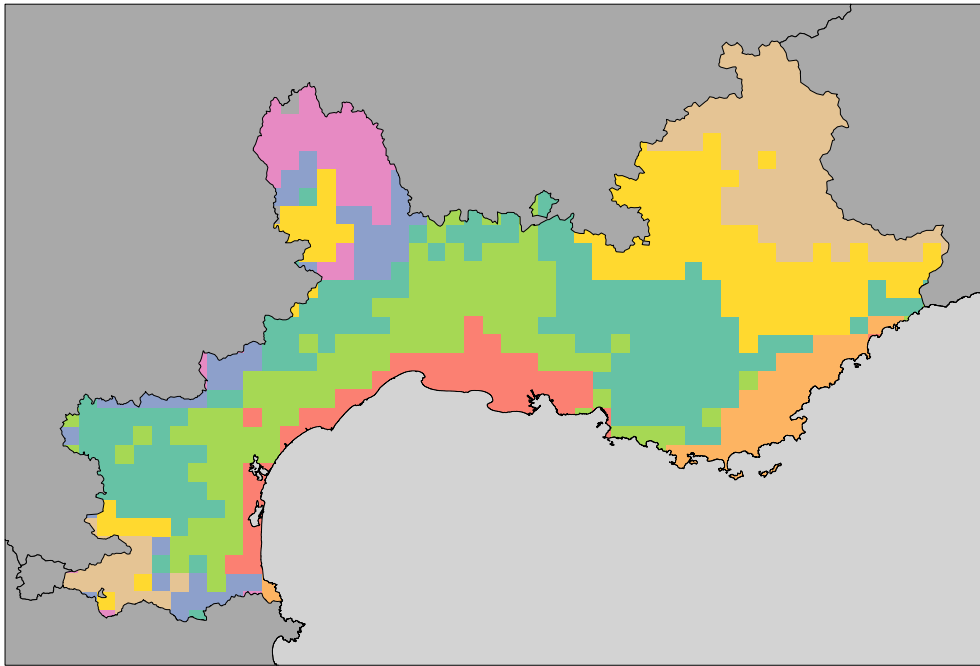


Figure S1. Biogeographical regions based on similarity in plant species ($l = 10$ km). Eight biogeographical regions have been identified. 1. Gulf of Lion coast in red. 2. Cork oak zone in orange. 3. Mediterranean lowlands in light green. 4. Mediterranean border in dark green. 5. Cévennes *sensu lato* in purple. 6. Subatlantic mountains in pink. 7. Prealps and other medium mountains in yellow. 8. High mountains in brown.

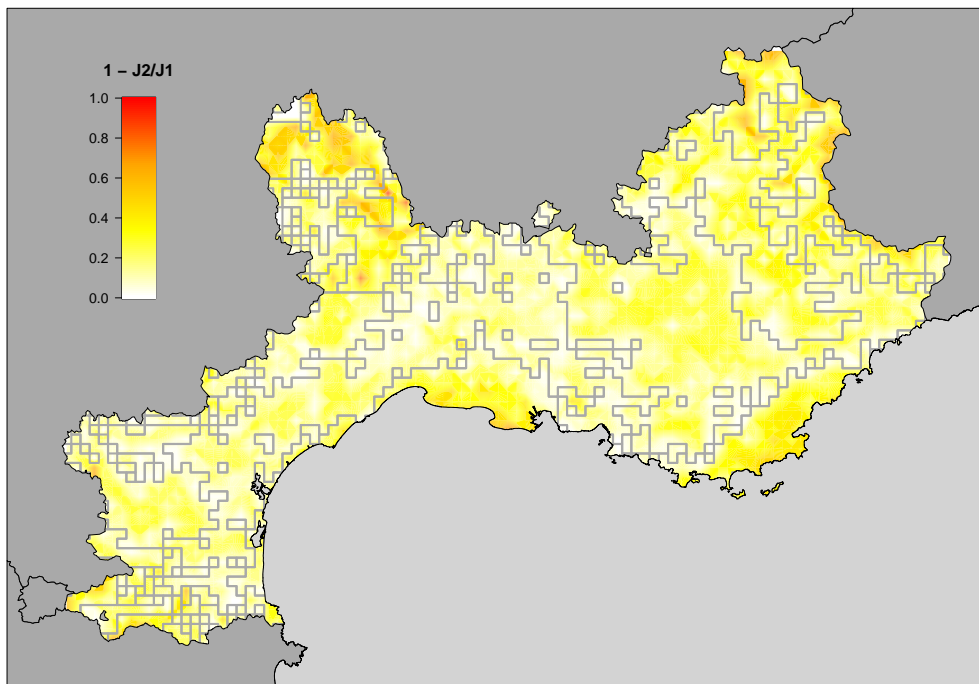


Figure S2. Uncertainty map ($l = 5$ km). For a given cell, $J1$ represents the average Jaccard similarity index between this cell and all the cells that belong to its cluster, and $J2$ represents the average Jaccard similarity index between this cell and all the cells belonging to the second closest cluster (based on the Jaccard similarity).

Table S1. Spatial coherence of the biogeographical regions according to the scale.

| Bioregion | n_i (l=5) | SP_i (l=5) | n_i (l=10) | SP_i (l=10) |
|-----------|-------------|--------------|--------------|---------------|
| 1 | 170 | 0.63 | 63 | 0.75 |
| 2 | 183 | 0.73 | 47 | 0.87 |
| 3 | 529 | 0.86 | 124 | 0.83 |
| 4 | 807 | 0.57 | 164 | 0.51 |
| 5 | 120 | 0.50 | 45 | 0.31 |
| 6 | 152 | 0.70 | 48 | 0.79 |
| 7 | 400 | 0.67 | 114 | 0.75 |
| 8 | 246 | 0.78 | 110 | 0.77 |

Table S2. Number of plant species per cluster of species.

| Cluster | Number of species |
|----------|-------------------|
| a | 445 |
| b | 149 |
| c | 230 |
| d | 299 |
| e | 169 |
| f | 277 |
| g | 37 |
| h | 180 |
| i | 242 |
| j | 136 |
| k | 125 |
| l | 95 |
| m | 180 |
| n | 180 |
| o | 178 |
| p | 186 |
| q | 44 |
| r | 59 |
| s | 212 |
| t | 274 |

Table S3. Network of interactions between biogeographical regions.

| Bioregion | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|-----------|-------------|-------------|-------------|-------------|------------|-------------|-------------|-------------|
| 1 | 0.52 | 0.24 | 0.15 | 0.07 | 0 | 0.01 | 0.01 | 0 |
| 2 | 0.15 | 0.57 | 0.1 | 0.14 | 0.01 | 0.01 | 0.02 | |
| 3 | 0.13 | 0.19 | 0.42 | 0.24 | 0 | 0 | 0.01 | 0 |
| 4 | 0.03 | 0.12 | 0.1 | 0.53 | 0.01 | 0 | 0.19 | 0.01 |
| 5 | 0.01 | 0.13 | 0 | 0.06 | 0.4 | 0.24 | 0.11 | 0.05 |
| 6 | 0.01 | 0.03 | 0 | 0.01 | 0.1 | 0.49 | 0.17 | 0.19 |
| 7 | 0 | 0.01 | 0 | 0.15 | 0.01 | 0.04 | 0.52 | 0.26 |
| 8 | 0 | 0 | 0 | 0.01 | 0.01 | 0.05 | 0.28 | 0.65 |