

# Towards a biogeographic regionalization of the European biota

Marta Rueda<sup>1\*</sup>, Miguel Á. Rodríguez<sup>1</sup> and Bradford A. Hawkins<sup>2</sup>

<sup>1</sup>Departamento de Ecología, Universidad de Alcalá, Alcalá de Henares, Madrid, Spain, <sup>2</sup>Department of Ecology & Evolutionary Biology, University of California, Irvine, CA, USA

## ABSTRACT

**Aim** To determine if it is possible to generate analytically derived regionalizations for multiple groups of European plants and animals and to explore potential influences on the regions for each taxonomic group.

Location Europe.

**Methods** We subjected range maps of trees, butterflies, reptiles, amphibians, birds and mammals to *k*-means clustering followed by *v*-fold cross-validation to determine the pattern and number of regions (clusters). We then used the mean range sizes of species in each group as a correlate of the number of regions obtained for each taxon, and climate and species richness gradients as correlates of the spatial arrangement of the group-specific regions. We also included the pattern of tree clusters as a predictor of animal clusters in order to test the 'habitat templet' concept as an explanation of animal distribution patterns.

**Results** Spatially coherent clusters were found for all groups. The number of regions ranged from three to eight and was strongly associated with the mean range sizes of the species in each taxon. The cluster patterns of all groups were associated with various combinations of climate, underlying species richness gradients and, in the case of animals, the arrangement of tree clusters, although the rankings of the correlates differed among groups. In four of five groups the tree pattern was the strongest single predictor of the animal cluster patterns.

**Main conclusions** Despite a long history of human disturbance and habitat modification, the European biota retains a discernable biogeographic structure. The primary driver appears to be aspects of climate related to water–energy balance, which also influence richness gradients. For many animals, the underlying habitat structure, as measured by tree distributions, appears to have a strong influence on their biogeographic structure, highlighting the need to preserve natural forest formations if we want to preserve the historical signal found in geographic distributions.

#### Keywords

Amphibians, biogeographic regions, birds, butterflies, Europe, habitat templet, *k*-means clustering, mammals, reptiles, trees.

\*Correspondence: Marta Rueda, Departamento de Ecología, Universidad de Alcalá, Ctra, Madrid-Barcelona, Km. 33.6, E-28871, Alcalá de Henares, Madrid, Spain. E-mail: marta.rueda@uah.es

## INTRODUCTION

How species are distributed in time and space has been a major research theme since the mid-19th century, when biogeographers began dividing the world into floral kingdoms and faunal regions using only their own knowledge of species distributions (Wallace, 1876; Engler, 1879-1882). More

© 2010 Blackwell Publishing Ltd

recently, the development of quantitative methods together with improved data availability has stimulated analytically derived biogeographic regionalizations at all levels of scaling (Birks, 1976; Myklestad & Birks, 1993; Williams *et al.*, 1999; Procheş, 2005; Moline & Linder, 2006; Heikinheimo *et al.*, 2007; Patten & Smith-Patten, 2008). Delineation of biogeographic regions is now frequently the initial step for conservation planning and management, and clustering procedures capable of capturing the spatial structure of species composition data are being used increasingly by conservation biologists as well as biogeographers (Proches, 2005; Finnie et al., 2007; Heikinheimo et al., 2007; Mackey et al., 2008; Patten & Smith-Patten, 2008). This approach has several advantages, including ease of comparison of the patterns across taxonomic groups, and the results are easily mapped (Mackey et al., 2008). Recently, Heikinheimo et al. (2007) applied cluster presence/absence methods to European land mammals and compared the resulting clusters against an environmental stratification of Europe based on climate, geomorphology and soil characteristics (see Metzger et al., 2005). They found that, despite a long history of human presence and habitat modification, Europe can be divided into cohesive subregions based on the distribution of mammals, consistent with their hypothesis that assemblages in the mammalian fossil record reflect ecologically distinct biogeographic units delimited by the environment (Heikinheimo et al., 2007).

The results of Heikinheimo et al. (2007) raise a number of interesting questions. For example, are biogeographic patterns within Europe taxon-specific? It is self-evident that phylogenetic and historical factors influence the taxonomic composition of an area, but such factors need not be identical everywhere, or for the entire biota. Given that species evolve within areas from which they may subsequently disperse, the taxonomic composition of an area's flora and fauna could reflect the degree to which it acts as a centre of origin, has been colonized by dispersing organisms, or has been subject to large-scale forces such as the Pleistocene glaciations (Mackey et al., 2008). Also, what drives the patterns, especially when we expect strong human influences on current plant and animal distributions? Extensive dispersion of plants and animals by humans over the past few hundred years, and the global extent of invasive species, represent other potential influences of species distributions. Finally, although climate can directly influence a region's taxonomic composition, it can also have indirect effects via impacts on geophysics and primary producers. For example, major soil groups - formed under different climatic conditions - influence plant species distributions. Plant distributions, in turn, influence faunal distributions (see Metzger et al., 2005, and references therein), generating a complex hierarchy of abiotic and biotic controls on the composition of biogeographic regions.

This latter effect can be directly connected to Southwood's (1977, 1988) concept of habitat templets, which argues that habitat provides the templet on which evolution forges animal life-history strategies. Southwood (1988; see also Mackey *et al.*, 2008) classified habitats with respect to three axes: (1) habitat productivity as a measure of potential growth, (2) habitat resource stability, and (3) adversity or stress. Over evolutionary time, these axes are presumed to operate as a selective force on the evolution of animal life histories, and on ecological time-scales the habitat templet filters out unsuccessful strategists from the pool of potential colonists, thereby controlling community composition (Scarsbrook & Townsend, 1993).

Given that vegetation structure and composition can be considered a habitat templet on which animal life-history strategies are shaped (Townsend *et al.*, 1997), we might expect regions based on plant distributions also to be reflected in animal-based regions.

Here, we evaluate biogeographic regionalizations for the trees, butterflies, amphibians, reptiles, birds and mammals of Europe. First, we use k-means clustering to define regions for each group. Second, we use the regions obtained for trees as a surrogate for habitat templet to evaluate the degrees of association with the regions found for animals. Trees should have strong influences on animal distributions, as they determine the vertical structure and horizontal patterning of vegetation, which are critical elements of animal habitat; further, trees are the dominant species in the most complex and stable ecosystems (Box, 1995). As an independent evaluation of the tree regionalization itself, we compare the regions with a map of natural vegetation in Europe (Bohn et al., 2000). We also use generalized linear models and variation partitioning techniques to identify the independent and shared influences of vegetation structure and climate of each taxonomic group on the distribution of biogeographic subregions within the continent. Finally, we explore associations between the clusters and the inherent species richness patterns of each group within our statistical framework. Our use of tree clusters to predict animal cluster analysis provides an example of how animal habitat fidelity can complement climatic-environmental approaches to understanding patterns of biogeographic regionalization.

## MATERIALS AND METHODS

#### **Distribution data**

Our database comprises all species of trees (187 species), butterflies (370 species), non-marine birds (342 species), nonmarine mammals (151 species), reptiles (78 species) and amphibians (59 species) that are native to continental western/ central Europe and Great Britain. Tree data were obtained from Montoya et al. (2007), who compiled and digitized tree species distribution maps from a variety of sources (for details see Montoya et al., 2007). Range maps were extracted from Cramp & Simmons (1977, 1980, 1983), Cramp (1985, 1988, 1992) and Cramp & Perrins (1993, 1994a,b) for birds; from Tolman & Lewington (2004) for butterflies; from the European Mammal Assessment database (IUCN, 2007, http://ec.europa.eu/environment/nature/conservation/species/ema/; accessed June 2008) for mammals; and from Arnold & Ovenden (2002) for reptiles and amphibians. It should be noted that these maps show the areas where a particular species can be expected to occur, although, within these areas, it will be found only in suitable habitats. Generally, although we refer to these maps as 'current distribution maps', data for the range maps were probably collected during the 19th and 20th centuries, although information has been updated as more knowledge about species distributions has been obtained. Accuracy of the maps can also vary – whereas for many countries up-to-date information on distributions is available, for others it may be scant, diffuse or hard to locate. In all cases, we accepted the maps at face value and digitized and rasterized them at *c*.  $50 \times 50$ -km grain size – based on the Universal Transverse Mercator (UTM) projection – to generate a presence/absence data grid. All islands except Great Britain, and all coastal cells containing < 50% of inland land mass, were excluded from the analysis, resulting in 1919 cells. We also calculated and mapped the species richness pattern of each taxonomic group.

Finally, we digitized the map of Bohn *et al.* (2000) of forestdominated formations to compare their distributions with those generated by our clustering procedure (see below). This map considers climate zones, ecological conditions and no human influence to identify nine major forest types across western and central Europe, based on expert opinion. We assigned these forest formations to our grid according to the dominant forest type occurring in each cell. The purpose of this analysis was to compare the *k*-means clusters against patterns for trees generated by an independent methodology to ensure our results are biologically meaningful.

#### **Environmental variables**

Three climatic variables were downloaded from http:// www.grid.unep.ch/data/summary.php?dataid (accessed October, 2008; data are representative of a time period extending approximately from 1920 to 1980). For each grid cell, we averaged annual precipitation (GNV174), mean annual temperature (GNV15) and Priestley-Taylor's annual actual evapotranspiration (AET) (GNV183) as measures of water availability, ambient energy and water-energy balance, respectively. Details and data sources for these variables are given by Ahn & Tateishi (1994a,b), Deichmann & Eklundh (1991) and Leemans & Wolfgang (1991). AET represents the annual balance between precipitation and latent heat exchange and is strongly associated with plant primary productivity (Rosenzweig, 1968). The three climatic variables selected are closely associated with species richness patterns of both plants and animals (O'Brien, 1993; Hawkins et al., 2003a,b; Rodríguez et al., 2005, 2006). See Fig. S1 in Supporting Information for the spatial patterns of the climatic variables.

#### **Clustering procedures**

We used *k*-means clustering (Han & Kamber, 2001; Bishop, 2005) combined with *v*-fold cross-validation (e.g. Molinaro *et al.*, 2005) to generate biotic regions for each taxon. We did not include spatial constraints during clustering; cells were clustered based on their species composition without regard to their spatial proximity (Legendre & Fortin, 1989, and references therein), so as to not force cohesion of clusters not justified by the actual distributions of species. The classical *k*-means clustering algorithm requires the number of clusters (k) to be established in advance, utilizes a subset of *k* random

initialization cells that are treated as the initial cluster centres, and then proceeds as a two-step iterative procedure in which cluster centres and clusters are successively recalculated. First, in the assignment step, each cell is assigned to its nearest cluster centre in terms of species compositional distance – we used Hellinger distances (for discussion and technical details, see Legendre & Gallagher, 2001; Gagné & Proulx, 2009). Secondly, in the update step, each cluster centre is updated by making it equal to the mean of the cells assigned to it. The process is repeated so that the clusters and cluster centres change in each replicate, and after a sufficient number of iterations (we used 100) they converge to a locally optimal position in the data space.

When k-means clustering is combined with v-fold crossvalidation, the potential number of k groups can be set as a range (we set these at two and 25 clusters), and the algorithm determines the 'best' number of clusters within these extremes given the data (see below). Commencing with the two-clusters case, this method divides the overall sample into randomly drawn subsamples of near-equal size, or  $\nu$  folds (we set  $\nu$  to 50, each comprising c. 40 cells). The cells belonging to the v - 1folds (c. 1869 cells) are then treated as a training sample and classified with k-means, and the clustering solution given by this analysis is applied to sample  $\nu$  to compute a measure of performance assessment or misclassification error (Molinaro et al., 2005). This consists of the average distance of cells to the cluster centres, and provides an estimate of how well the cells in sample  $\nu$  can be assigned to the cluster solution. The process is replicated until all 50 folds are cross-examined, and the results are averaged to yield a single misclassification error estimate for the two-clusters case. The analysis is repeated for three clusters, four clusters, and so on, until a (k + 1)-clusters solution gives an average misclassification error that is almost equal to that obtained for k clusters – we set a 5% threshold level of error disparity, that is, solutions differing less than this threshold were considered equal in terms of misclassification error levels. When this occurs, the k number of clusters is identified as the 'best' number and used to generate a k-means classification utilizing all data. Finally, k-means results depend to some extent on the initialization cells that are used to represent the initial cluster centres (Heikinheimo et al., 2007). To minimize this dependence, we ran the entire classification process 100 times using a different set of randomly chosen initialization cells in each run, and we retained the solution with the least error as the final solution (Heikinheimo et al., 2007). The k-means clustering and v-fold cross-validation were carried out using the data-mining module in STATISTICA 8.0 (StatSoft, Tulsa, OK, 2007).

#### Generalized linear models (GLMs) of k-means clusters

If performed with a generalized multinomial logit-link, GLMs can be used for modelling a multinomial response variable (e.g. the k-means cluster solution obtained for a particular taxon) as a function of one or more continuous and/or categorical predictors. We generated single- and multiple-

variable explanatory models of our biogeographic regionalization results using the Generalized Linear Model module in STATISTICA. Predictors included the climate variables individually and in concert, the species richness of the taxon, and, for the animal taxa, the clusters obtained for trees, a multinomial variable used to evaluate the extent to which animal species composition is associated with tree species composition. We also used GLM to analyse the similarity between our tree clusters and the map of forest-dominated formations (Bohn *et al.*, 2000). In all cases, the strengths of the relationships were estimated as the proportion of the deviance in the response variable that was described by the model.

It should be noted that, for amphibians and reptiles, the groups for which we obtained the most clusters (see below), results from complex GLMs that included the tree groups as a predictor could be generated only if the 0 s in their contingency tables with the tree groups were changed to 1 s (we reclassified 21 cells in the case of reptiles and 18 cells for amphibians). These changes were made by randomly selecting the cells that were modified from the sets of potential candidate cells. The total deviances of the clusterings of both groups, which were > 7000 units in both cases, were changed by less than one unit, so the small modifications made to run the program did not alter the final results. Finally, we used deviance partitioning (Lobo et al., 2002) to partition the deviance of each response variable into independent 'effects' of a particular predictor and covarying 'effects' of two or more predictors that cannot be disentangled.

## Geographic range sizes and animal clusters

The number of clusters obtained for any taxonomic group should be influenced by the range size distribution on which the clusters are based, because a single cluster would be obtained if all species were distributed over the entire region, whereas many clusters would potentially be identified if all species were narrowly distributed endemics in different parts of the region. We evaluated this by rank correlating the number of clusters found for each group against the mean range size of that taxon, measured as the arithmetic mean number of grid cells occupied by each species (maximum = 1919).

## RESULTS

## Tree clusters and European forest formations

Although we did not include spatial constraints to influence the clustering of neighbouring cells, biogeographic subregions for trees generated by k-means clustering were strongly cohesive spatially, and moderately coincident with the distributions of the main natural forest formations of Europe recognized by Bohn *et al.* (2000) (Fig. 1), which explained 49.7% of the deviance in the k-means clusters. In general, k-means clustering was unable to distinguish smaller-scale forest formations (e.g. the alpine forests of the Alps and Pyrenees) or the subarctic vs. alpine forests of Scandinavia (cf. Fig. 1a,b). Further insensitivity of *k*-means to small-scale patterns is notable across southern Europe (Fig. 1b). Even so, it is possible to match the Iberian 'red' cluster with 'Mediter-ranean sclerophyllous' forests (Fig. 1); the southern 'orange' cluster loosely identifies 'thermophilous mixed deciduous broadleaved' forests; the central 'green' cluster approximates the distribution of 'mesophytic deciduous broadleaved and coniferous–broadleaved forests'; and the Scandinavian/Scottish 'dark blue' cluster coincides with 'mesophytic and hygromesophytic coniferous and broadleaved–coniferous' forests. The south-eastern 'yellow' cluster (Fig. 1b) does not match any forest formation and might be interpreted as a heterogeneous region with a mixture of forest types (Fig. 1a).

All in all, the k-means clustering of trees captured what are likely to represent meaningful biological units rather than entirely arbitrary constructs generated by the clustering algorithm, albeit with some loss of resolution (five forest groups rather than the seven main types recognized by Bohn *et al.*, 2000) and no small-scale detail, which may be due to the relatively coarse grain of our grid. Therefore, we treat the tree clustering pattern as the spatial configurations of species constellations that can be compared with the cluster patterns of the animal taxa as a habitat templet.

## Animal clusters

As with trees, we obtained cohesive subregions for all animal groups (Fig. 2). However, there was substantial variation in the number of clusters for each group, ranging from three for birds to eight for reptiles. Also, although there were similarities in some cluster borders for the various groups (Fig. 2), none of the clustering patterns was identical. The extent of congruence of clusters with tree clusters was also variable among the animal groups (cf. Figs 1 & 2). Across all groups (including trees), the number of clusters is strongly associated with the mean range sizes of the constituent species, as expected (Fig. 3).

## **Correlates of clustering patterns**

In single-factor models, climate explained more than half of the deviance in the cluster patterns of plants and animals in all cases (Table 1), and models combining temperature, precipitation and AET always explained substantially more deviance than did models containing the individual components of climate. In five groups, the three-factor models were the best models according to the Akaike information criterion  $(\Delta AIC > 27$  for all other models), whereas for birds the three-factor model and a two-factor model including AET and temperature were statistically indistinguishable ( $\Delta AIC = 0.5$ ) and explained similar levels of deviance (66.0% vs. 65.9%). Within the climate models, AET explained more deviance than either temperature or precipitation in all groups except birds (Table 1), identifying water-energy balance as a strong correlate of both plant and animal distribution patterns across Europe. The extent to which the underlying richness patterns



Figure 1 (a) Native forest formations in Europe distinguished by Bohn et al. (2000); (b) k-means clusters after v-fold cross-validation.

also reflect clustering patterns ranged from minor for birds to substantial for reptiles (Table 1; see Fig. S2 for the species richness patterns of each group). Tree clusters explained more deviance than climate or species richness in all groups except reptiles. However, the deviance explained by tree clusters was considerably larger for birds and butterflies. Finally, we correlated species richness among taxa to evaluate the extent that the similarity of paired cluster patterns is associated with the underlying similarity of richness gradients (Table 2). As there was a moderate correlation between the deviance values among combinations of groups and their correlations in richness (r = 0.551, P < 0.05), the similarity of cluster patterns partially depends on the similarity of richness patterns, or on the environmental drivers underlying richness.

The partitions of deviance indicated that the largest fraction of the variability in animal clusters was accounted for by the combined 'effect' of the three groups of explanatory variables (climate, species richness and tree clusters) for all groups except birds, for which the largest fraction (47.6%) was explained by the joint 'effect' of climate and tree clusters (Fig. 4). The variation attributable to the independent 'effect' of tree clusters was higher than that of the independent 'effect' of climate or species richness for all animal taxa except reptiles, and ranged from 17.6% for mammals to 21.8% for butterflies. For reptiles, the independent 'effect' of the species richness pattern (18.6%) exceeded the independent 'effect' of tree clusters (10.7%), whereas it was negligible for birds and mammals, slight for butterflies (6.8%) and moderate for amphibians (10.4%). The independent 'effect' of climate was negligible for butterflies and reptiles and moderate for birds, mammals and amphibians (9.4%, 8.9% and 12.4%, respectively). Thus the multivariable models indicate that, in most cases, the inherent covariance among climate, vegetation and species richness gradients makes it difficult to identify unambiguously the primary correlate of the cluster patterns, but for most animal groups, tree distributions do matter.

## DISCUSSION

First and foremost, the claim by Heikinheimo *et al.* (2007) that Europe retains spatially cohesive biogeographic subregions based on the distributions of mammals holds for all plant and animal taxa we examined. On the other hand, we find substantial variation in the number of 'regions' for each group (from three for birds to eight for reptiles), largely dependent on how widely species in each group are distributed. The cluster



**Figure 2** The k-means clustering of the c. 50-km cells in Europe for (a) birds, (b) butterflies, (c) mammals, (d) reptiles, (e) amphibians. The result in each case is the best of 100 clustering runs in terms of misclassification error.

patterns also differ in their detailed spatial arrangements, indicating that, as a whole, western and central Europe does not have a simple biogeographic structure. Rather, regionalization patterns depend on the taxonomic group of interest. The structures of our explanatory models suggest that the main factors influencing biogeographic units also differ among taxa. However, the strong deviance partitioned when confronting animal clusters with tree clusters indicates that habitat, as defined by constellations of tree species, is probably



**Figure 3** Relationship between the number of *k*-means clusters and the arithmetic mean range size ( $\pm$  1 SE) of the species in each European taxonomic group studied. Range sizes are estimated as the number of grid cells occupied by each species. The Spearman rank correlation is -0.986 (P < 0.001).

**Table 1** Percentage of deviance in each European taxon's k-means groups explained by mean annual temperature (T), annual precipitation (P), annual actual evapotranspiration (AET), the climatic variables combined, the taxon's species richness pattern, and the k-means clusters obtained for trees.

Taxon	Т	Р	AET	Climate $(P + T + AET)$	Species richness	Tree clusters
Trees	34.0	3.1	39.5	57.0	48.9	_
Birds	47.6	0.5	47.1	66.0	9.3	74.1
Butterflies	26.6	1.9	46.5	58.1	50.4	73.2
Mammals	36.5	1.1	47.5	64.2	41.9	68.2
Reptiles	27.7	2.9	35.4	50.2	69.6	51.5
Amphibians	23.1	4.5	38.7	52.7	41.6	55.1

**Table 2** Associations among paired European taxa between

 cluster patterns (measured by percentage deviance, below diagonal) and species richness patterns (correlation coefficients, above diagonal).

	Trees	Birds	Butterflies	Mammals	Reptiles	Amphibians
Trees	_	0.420	0.895	0.887	0.689	0.735
Birds	74.1	-	0.528	0.506	0.279	0.454
Butterflies	73.3	46.5	-	0.924	0.657	0.740
Mammals	68.2	54.7	69.7	-	0.595	0.821
Reptiles	51.5	36.7	50.0	49.8	_	0.503
Amphibians	55.1	40.9	56.9	50.7	59.3	-

an important driver of animal distributions, except perhaps for those groups that are more constrained by particular life history traits, such as reptiles. The largest fraction of variability is accounted for by the combined 'effect' of species richness patterns, climate and tree clusters, indicating that true independent associations of potential abiotic and biotic drivers cannot be identified unambiguously. Irrespective, it appears



**Figure 4** Results of the deviance partitioning analysis for (a) trees, (b) birds, (c) butterflies, (d) mammals, (e) reptiles, (f) amphibians in Europe, as explained by climate (mean annual temperature + annual precipitation + annual actual evapotranspiration); species richness of the group; and for animals (b–f) k-means groups obtained for trees. Deviance fractions < 5% are indicated in white; between 5–15%, light grey; > 15%, dark grey.

that, despite the long history of human presence and subsequent habitat transformation and disturbance that has occurred in Europe over the past several millennia, animal distributions remain related to natural factors such as topography and climate, as suggested by Heikinheimo *et al.* (2007), and continue to reflect a relatively strong influence of animals' habitat fidelities. That is, despite wholesale modification of natural habitats over large geographic areas by humans, the habitat templet concept (*sensu* Southwood, 1977, 1988) remains a useful tool for understanding animal distribution patterns at the continental spatial scale.

We also find differences among animal groups that suggest linkages between biological traits and patterns of regionalization. For example, the larger number of clusters for amphibians and reptiles clearly reflects that most species are narrowly distributed, which, in part, is probably explained by the limited dispersal of both taxa compared with the high vagility of birds and butterflies (Bowne & Bowers, 2004). Amphibians and reptiles also share basic underlying traits that distinguish them from the other vertebrate groups; they are ectotherms. Whereas reptiles are extreme solar ectotherms, amphibians, although ectothermic and susceptible to temperature as well, usually require moisture for survival and water for reproduction. These physiological, morphological and life-history traits probably make amphibians and reptiles more susceptible to habitat idiosyncrasies, which then lead to narrower ranges and more strongly regionalized faunas. Future analyses in other regions will determine if the similarities and differences we observe for the cluster patterns among groups are driven by biology, or by the particular history of Europe.

Lastly, there are differences among groups with respect to the covariation between cluster patterns and underlying patterns of species richness. Birds, mammals and, to a lesser, extent butterflies cluster more or less independently of their richness gradients, whereas amphibians and reptiles show stronger covariation. Reptiles and amphibians contain the fewest number of species in comparison with the other groups, which probably gives rare species more weight when clustering. Even if true, regions are not necessarily defined around highdiversity areas, and biogeographic structure at the continental scale is not simply a reclassification of species richness gradients.

Although the focus of our analysis was not on conservation, and our maps should not be used to generate conservation plans for any major taxonomic group, the regionalization patterns within Europe clearly show the importance of habitat structure, especially forests, to the maintenance of broad-scale biogeographic patterns of animals, as also argued by Escalante *et al.* (2007) for Mexico. We can conclude that, although it is still possible to track evolutionary and ecological processes when exploring the present-day distributions of species in Europe, it becomes a priority to arrest habitat loss if we want to conserve not only biodiversity, but also the ecological and evolutionary signals that underlie species richness patterns.

## ACKNOWLEDGEMENTS

This study was supported by the Spanish Ministry of Science and Innovation (grants PR2009-0107 and CGL2006-03000/ BOS to M.Á.R.). M.R. was supported with a postdoctoral grant from the University of Alcalá and thanks the University of California Irvine for their hospitality during the preparation of this manuscript. We also thank Ole Vetaas and two anonymous referees for their help in improving the manuscript. This work is part of the 'Current and past determinants of Western Palaearctic biodiversity' working group, organized by Joaquín Hortal and funded by the NERC Centre for Population Biology Workshop series (2008–10).

## REFERENCES

Ahn, C.H. & Tateishi, R. (1994a) Development of a global 30-minute grid potential evapotranspiration data set. *Jour-*

nal of the Japan Society of Photogrammetry and Remote Sensing, **33**, 12–21.

- Ahn, C.H. & Tateishi, R. (1994b) Development of global evapotranspiration and water balance data sets. *Journal of the Japan Society of Photogrammetry and Remote Sensing*, **33**, 48–61.
- Arnold, E.N. & Ovenden, D. (2002) A field guide to the reptiles and amphibians of Britain and Europe. Harper Collins, London.
- Birks, H.J.B. (1976) The distribution of European pteridophytes: a numerical analysis. *New Phytologist*, **77**, 257–287.
- Bishop, C.M. (2005) Neural networks for pattern recognition. Oxford University Press, New York.
- Bohn, U., Gollug, G. & Hettwer, C. (2000) *Map of the natural vegetation of Europe*. Federal Agency for Nature Conservation, Bonn.
- Bowne, D.R. & Bowers, M.A. (2004) Interpatch movements in spatially structured populations: a literature review. *Landscape Ecology*, **19**, 1–20.
- Box, E.O. (1995) Factors determining distributions of tree species and plant functional types. *Vegetation*, **121**, 101–116.
- Cramp, S. (ed.) (1985) Handbook of the birds of Europe the Middle East and North Africa. The birds of the Western Palearctic. Vol. IV: Terns to woodpeckers. Oxford University Press, Oxford.
- Cramp, S. (ed.) (1988) Handbook of the birds of Europe the Middle East and North Africa. The birds of the Western Palearctic. Vol. V: Tyrant flycatchers to thrushes. Oxford University Press, Oxford.
- Cramp, S. (ed.) (1992) Handbook of the birds of Europe the Middle East and North Africa. The birds of the Western Palearctic. Vol. VI: Warblers. Oxford University Press, Oxford.
- Cramp, S. & Perrins, C.M. (eds) (1993) Handbook of the birds of Europe the Middle East and North Africa. The birds of the Western Palearctic. Vol. VII: Flycatchers to shrikes. Oxford University Press, Oxford.
- Cramp, S. & Perrins, C.M. (eds) (1994a) Handbook of the birds of Europe the Middle East and North Africa. The birds of the Western Palearctic. Vol. VIII: Crows to finches. Oxford University Press, Oxford.
- Cramp, S. & Perrins, C.M. (eds) (1994b) Handbook of the birds of Europe the Middle East and North Africa. The birds of the Western Palearctic. Vol. IX: Buntings and New World warblers. Oxford University Press, Oxford.
- Cramp, S. & Simmons, K.E.L. (eds) (1977) Handbook of the birds of Europe the Middle East and North Africa. The birds of the Western Palearctic. Vol. I: Ostrich to ducks. Oxford University Press, Oxford.
- Cramp, S. & Simmons, K.E.L. (eds) (1980) Handbook of the birds of Europe, the Middle East and North Africa. The birds of the Western Palearctic. Vol. II: Hawks to bustards. Oxford University Press, Oxford.
- Cramp, S. & Simmons, K.E.L. (eds) (1983) Handbook of the birds of Europe, the Middle East and North Africa. The birds

of the Western Palearctic. Vol. III: Waders to gulls. Oxford University Press, Oxford.

- Deichmann, U. & Eklundh, L. (1991) *Global digital datasets for land degradation studies: a GIS approach.* UNEP, Nairobi, Kenya.
- Engler, A. (1879-1882) Versuch einer Entwicklungsgeschichte der *Pflanzenwelt*. Engelman, Leipzig.
- Escalante, T., Sánchez-Cordero, V., Morrone, J.J. & Linaje, M. (2007) Deforestation affects biogeographical regionalization: a case study contrasting potential and extant distributions of Mexican terrestrial mammals. *Journal of Natural History*, **41**, 965–984.
- Finnie, T.J.R., Preston, C.D., Hill, M.O., Uotila, P. & Crawley, J.C. (2007) Floristic elements in European vascular plants: an analysis based on *Atlas Florae Europaeae*. *Journal of Biogeography*, **34**, 1848–1872.
- Gagné, S. & Proulx, R. (2009) Accurate delineation of biogeographical regions depends on the use of an appropriate distance measure. *Journal of Biogeography*, **36**, 561–567.
- Han, J. & Kamber, M. (2001) Cluster analysis. *Data mining: concepts and techniques* (ed. by D.D. Cerra, H. Severson and B. Breyer), pp. 349–351. Morgan Kaufmann, San Francisco, CA.
- Hawkins, B.A., Field, R., Cornell, H.V., Currie, D.J., Guégan, J.F., Kaufman, D.M., Kerr, J.K., Mittelbach, G.G., Oberdorff, T., O'Brien, E.M., Porter, E.E. & Turner, J.R.G. (2003a) Energy, water, and broad-scale geographic patterns of species richness. *Ecology*, 84, 3105–3117.
- Hawkins, B.A., Porter, E.E. & Diniz-Filho, J.A.F. (2003b) Productivity and history as predictors of the latitudinal diversity gradient of terrestrial birds. *Ecology*, 84, 1608–1623.
- 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.
- Leemans, R. & Wolfgang, P.C. (1991) The IIASA Database for mean monthly values of temperature, precipitation and cloudiness of a global terrestrial grid. Report no. RR-91-018. International Institute for Applied Systems Analysis (IIA-SA), Laxenburg, Austria.
- Legendre, P. & Fortin, M.J. (1989) Spatial pattern and ecological analysis. *Plant Ecology*, **80**, 107–138.
- Legendre, P. & Gallagher, E.D. (2001) Ecologically meaningful transformations for ordination of species data. *Oecologia*, **129**, 271–280.
- Lobo, J.M., Lumaret, J.P. & Jay-Robert, P. (2002) Modelling the species richness distribution of French dung beetles (Coleoptera, Scarabaeidae) and delimiting the predictive capacity of different groups of explanatory variables. *Global Ecology and Biogeography*, **11**, 265–277.
- Mackey, B.G., Berry, S.L. & Brown, T. (2008) Reconciling approaches to biogeographic regionalization: a systematic and generic framework examined with a case study of the Australian continent. *Journal of Biogeography*, **35**, 213–229.
- Metzger, M.J., Bunce, R.G.H., Jongman, R.H.G., Mücher, C.A. & Watkins, J.W. (2005) A climatic stratification of the

environment of Europe. *Global Ecology and Biogeography*, **14**, 549–563.

- Molinaro, A., Simon, R. & Pfeiffer, R.M. (2005) Prediction error estimation: a comparison of resampling methods. *Bioinformatics*, **21**, 3301–3307.
- Moline, P.M. & Linder, H.P. (2006) Input data, analytical methods and biogeography of *Elegia* (Restionaceae). *Journal of Biogeography*, **33**, 47–62.
- Montoya, D., Rodríguez, M.Á., Zavala, M.Á. & Hawkins, B.A. (2007) Contemporary richness of Holarctic trees and the historical pattern of glacial retreat. *Ecography*, **30**, 173–182.
- Myklestad, Å. & Birks, H.J.B. (1993) A numerical analysis of the distribution patterns of *Salix* L. species in Europe. *Journal of Biogeography*, **20**, 1–32.
- O'Brien, E.M. (1993) Climatic gradients in woody plant species richness: towards an explanation based on an analysis of southern Africa's woody flora. *Journal of Biogeography*, **20**, 181–198.
- Patten, M.A. & Smith-Patten, B.D. (2008) Biogeographical boundaries and Monmonier's algorithm: a case study in the northern Neotropics. *Journal of Biogeography*, 35, 407–416.
- Procheş, Ş. (2005) The world's biogeographical regions: cluster analyses based on bat distributions. *Journal of Biogeography*, 32, 607–614.
- Rodríguez, M.Á., Belmontes, J.A. & Hawkins, B.A. (2005) Energy, water and large-scale patterns of reptile and amphibian species richness in Europe. *Acta Oecologica*, **28**, 65–70.
- Rodríguez, M.Á., López-Sañudo, I.L. & Hawkins, B.A. (2006) The geographic distribution of mammal body size in Europe. *Global Ecology and Biogeography*, **15**, 185–193.
- Rosenzweig, M.L. (1968) Net primary productivity of terrestrial environments: predictions from climatological data. *The American Naturalist*, **102**, 67–74.
- Scarsbrook, M.R. & Townsend, C.R. (1993) Stream community structure in relation to spatial and temporal variation: a habitat templet study of two contrasting New Zealand streams. *Freshwater Biology*, **29**, 395–410.
- Southwood, T.R.E. (1977) Habitat, the templet for ecological strategies? *Journal of Animal Ecology*, **46**, 337–365.
- Southwood, T.R.E. (1988) Tactics, strategies and templets. *Oikos*, **52**, 3–18.
- Tolman, T. & Lewington, R. (2004) *Collins field guide. The butterflies of Britain and Europe*, 2nd edn. Harper Collins, London.
- Townsend, C.R., Doledec, S. & Scarsbrook, M.R. (1997) Species traits in relation to temporal and spatial heterogeneity in streams: a test of habitat templet theory. *Freshwater Biology*, **37**, 367–387.
- Wallace, A.R. (1876) *The geographical distribution of animals*. Macmillan, London.
- Williams, P.H., de Klerk, H.M. & Crowe, T.H. (1999) Interpreting biogeographical boundaries among Afrotropical birds: spatial patterns in richness gradients and species replacement. *Journal of Biogeography*, 26, 459–474.

M. Rueda et al.

## SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

**Figure S1** Spatial patterns for (a) annual actual evapotranspiration (following Priestley–Taylor's formula) (mm); (b) annual mean temperature (°C); (c) annual precipitation (mm) in Europe. These variables are available at http:// www.grid.unep.ch/data/summary.php?dataid (accessed October 2008).

**Figure S2** Patterns of species richness for (a) trees, (b) birds, (c) butterflies, (d) mammals, (e) reptiles and (f) amphibians in Europe.

As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials are peer-reviewed and may be re-organized for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.

## BIOSKETCHES

**Marta Rueda** is interested in biogeography and conservation biology. Currently she studies how habitat loss and fragmentation at broad scales influence extinction, especially of forest birds. Her main research interests also include aspects of plant–animal interactions at the community scale.

**Miguel Á. Rodríguez** is interested in the processes conditioning patterns of biodiversity at regional and global scales. His recent research has involved investigating the effects of habitat destruction and fragmentation on aggregate properties of faunas at broad scales.

**Bradford A. Hawkins** is interested in biogeography and geographical ecology, with an emphasis on the ecological and evolutionary influences on species richness gradients.

Editor: Ole Vetaas