**RESEARCH PAPER**

# Patterns in species richness and endemism of European freshwater fish

Yorick Reyjol[1]*, Bernard Hugueny[1], Didier Pont[1]†, Pier Giorgio Bianco[2], Ulrika Beier[3], Nuño Caiola[4], Frederic Casals[5], Ian Cowx[6], Alcibiades Economou[7], Teresa Ferreira[8], Gertrud Haidvogl[9], Richard Noble[6], Adolfo de Sostoa[4], Thibault Vigneron[10] and Tomas Virbickas[11]

[1]*Laboratory of Fluvial Hydrosystems Ecology, University Lyon I, France,* [2]*Department of Biology, University of Napoli, Italy,* [3]*Institute of Freshwater Research, Drottningholm, Sweden,* [4]*Department of Animal Biology — Vertebrates, University of Barcelona, Spain,* [5]*Department of Animal Production — Wildlife, University of Lleida, Spain,* [6]*International Fisheries Institute, University of Hull, UK,* [7]*Hellenic Centre for Marine Research, Athens, Greece,* [8]*Forest Department, Technical University of Lisbon, Portugal,* [9]*University of Natural Resources and Applied Life Sciences, Vienna, Austria,* [10]*Fisheries French Council, Cesson-Sévigné, France, and* [11]*Laboratory of Hydrobiont Ecology and Physiology, University of Vilnius, Lithuania*

*Correspondence: Yorick Reyjol, Département de Chimie-Biologie, Université du Québec à Trois-Rivières, Québec Canada. E-mail: yorick.reyjol@uqtr.ca
†Present address: Cemagref, Hydrobiological Research unit, Aix en Provence, France.*

## ABSTRACT

**Aim** To analyse the patterns in species richness and endemism of the native European riverine fish fauna, in the light of the Messinian salinity crisis and the Last Glacial Maximum (LGM).

**Location** European continent.

**Methods** After gathering native fish faunistic lists of 406 hydrographical networks, we defined large biogeographical regions with homogenous fish fauna, based on a hierarchical cluster analysis. Then we analysed and compared the patterns in species richness and endemism among these regions, as well as species–area relationships.

**Results** Among the 233 native species present in the data set, the Cyprinidae family was strongly dominant (> 50% of the total number of species). Seven biogeographical regions were defined: Western Peri-Mediterranea, Central Peri-Mediterranea, Eastern Peri-Mediterranea, Ponto-Caspian Europe, Northern Europe, Central Europe and Western Europe. The highest regional species richness was observed for Central Peri-Mediterranea and Ponto-Caspian Europe. The highest endemic richness was found in Central Peri-Mediterranea. Species–area relationships were characterized by high slope values for Peri-Mediterranean Europe and low values for Central and Western Europe.

**Main conclusions** The results were in agreement with the 'Lago Mare' hypothesis explaining the specificity of Peri-Mediterranean fish fauna, as well as with the history of recolonization of Central and Western Europe from Ponto-Caspian Europe following the LGM. The results also agreed with the mechanisms of speciation and extinction influencing fish diversity in hydrographical networks. We advise the use of the seven biogeographical regions for further studies, and suggest considering Peri-Mediterranean Europe and Ponto-Caspian Europe as 'biodiversity hotspots' for European riverine fish.

**Keywords**
Biodiversity hotspots, biogeographical regions, endemism, Europe, hydrographical network, Messinian salinity crisis, native fish fauna, Pleistocene glaciations, species–area relationships, species richness.

## INTRODUCTION

There is evidence that patterns in species richness and endemism can be related to historical events (Banarescu, 1990, 1992; Matthews, 1998; Hewitt, 1999, 2000). During the Quaternary era, i.e. from 2.4 million years ago (Ma) until now, the Earth has been considerably affected by cyclic glacial events, notably related to variations of the Earth's orbit around the sun (the

Croll–Milankovich theory; Bennett, 1997; Williams *et al.*, 1998). These thermal cycles have led to successive contractions and extensions of the geographical ranges of species, for both plants and animals (Blondel & Vigne, 1993; Hewitt, 1999, 2000).

In Europe during the Last Glacial Maximum (LGM; from 24,000 to 18,000 years ago), the ice cap separated the continent into two parts, the northern part being covered either by ice or by permafrost, while the southern part (mainly areas surrounding

the Mediterranean Sea, i.e. Peri-Mediterranean areas) was spared (see Hewitt, 2000 for a map). As a consequence, numerous terrestrial and aerial species took shelter in Peri-Mediterranean refuges (i.e. the Balkans, the Caucasus and the Iberian and Italian peninsulas), from which Europe was recolonized after the LGM, following a 'hedgehog', 'bear' or 'grasshopper' pattern (Hewitt, 1999, 2000). Hewitt (2000) suggested that these patterns could be used as paradigms, but also stressed the need for other studies to make conclusions. For instance, Bilton *et al.* (1998) showed that for small mammals, Peri-Mediterranean Europe was more an area of endemism (by allopatric speciation) than a source for northwards post-glacial recolonization. On the contrary, they showed that recolonization may have occurred from refuges in Central Europe and Western Asia (Soffer, 1990).

Among freshwater fish, primary and primary-like forms such as the ostariophysians (*sensu* Myers, 1951), inclusive of about 80% of all freshwater fish, are unable to disperse through sea waters and are therefore restricted to the hydrographical networks of drainage basins (i.e. the land area where precipitation runs off into rivers and lakes). Therefore, in the absence of human intervention, their dispersal relies entirely on the geomorphological evolution of hydrographical networks. Colonization between basins can only take place on a long temporal scale, i.e. during marine regressions when sea level decreases and downstream connections between basins become possible, or during orogenesis, which allows river captures between opposite sides of mountains (Conner & Suttkus, 1986; Banarescu, 1990, 1992; Bianco, 1995). For instance, the ability of cold water species to 'cross' high mountains was recently pointed out for the bullhead, *Cottus gobio* L., which occurs on both sides of the Alps (Slechtova *et al.*, 2004). As a consequence, it is usually supposed that, among vertebrates, while numerous birds and mammals migrated towards Peri-Mediterranean areas during the LGM (Blondel & Vigne, 1993; Hewitt, 1999, 2000), many fish species were unable to migrate along a north–south axis during this time, therefore becoming extinct (Moyle & Herbold, 1987; Banarescu, 1990, 1992). The main surviving species were those present in the Danube Basin before the glaciations, whose fauna derived from the freshwater or oligosaline phase of the Parathetys, during the Miocene (for details see Banarescu, 1990, 1992; Bianco, 1990). The colonization of Peri-Mediterranean areas by fish occurred a long time before the LGM (Bianco, 1990), when fresh water from Parathetys drained into and filled the dry or nearly dried up Mediterranean Sea, immediately after the Messinian salinity crisis (about 5 Ma) (Hsü *et al.*, 1977).

The aim of the present work was to analyse patterns in species richness and endemism for the European riverine fish fauna in the light of the two major events that profoundly affected fish extinction, dispersal and speciation since the end of the Miocene: the Messinian salinity crisis of the Mediterranean Sea (5 Ma) and the LGM (18,000 years ago). To do so, we first defined large biogeographical regions with homogeneous native fauna, using lists of fish fauna from hydrographical basins distributed throughout Europe. Indeed, no single factor is more important in the regional biogeography of freshwater fish than drainage basin limits (Gilbert, 1980), and hydrographical networks can be

considered as biogeographical islands, containing specific pools of species (Livingstone *et al.*, 1982; Hugueny, 1989a,b). Then, we compared the number of species and the level of endemism on the biogeographical scale. Moreover, as Westoby (1993) pointed out, comparisons of species–area relationships between biogeographical provinces offer the only available means of determining whether evolutionary history has an important influence on extant diversity, so species–area relationships were compared among biogeographical regions.

## MATERIALS AND METHODS

### Data set

Faunistic lists from 406 basins distributed throughout Europe, i.e. from England and the Iberian Peninsula in the west (longitude 10° W) to the Ural mountains in the east (longitude 60° E) were collected from both the published and the grey literature. No data were collected for Iceland, Ireland or Norway. Only native species per basin were considered, based both on expert knowledge and the literature, including migratory species but excluding obligatorily estuarine species with no freshwater life stage. The different forms of the brown trout, i.e. *Salmo trutta fario* L., *Salmo trutta lacustris* L. and *Salmo trutta trutta* L., were not distinguished, and all *Coregonus* species were excluded because of some taxonomical uncertainties. A total of 233 species were present in the data set.

### Definition of biogeographical regions

#### Regrouping of small basins

Preliminary hierarchical clustering on the $406 \times 233$ presence–absence matrix showed that considering each drainage basin independently led to uninterpretable results with regard to geographical coherence. This notably originated from the increase in species richness with surface area of the drainage basin, as previously shown by Livingstone *et al.* (1982) and Hugueny (1989a,b). Therefore, to limit the influence of basin size in hierarchical clustering, only large drainage basins (surface area > 25,000 km$^2$) were considered independently for the present work. Smaller basins were grouped according to the marine area they drain to, considering that neighbouring basins could have potentially been colonized by primary or primary-like freshwater fish species during sea regressions (Conner & Suttkus, 1986; Banarescu, 1990, 1992; Bianco, 1995). The marine areas used are the ICES Fishing Areas (http://www.ices.dk), or the FAO marine areas (http://www.fao.org). The fishing areas which comprised peninsular or insular coasts of Europe (i.e. the English Channel, and the Baltic, Ionian, Adriatic and North seas) were separated in two, because they may have undergone different biogeographical histories (Table 1). Since only one basin was available for the Spanish coast of the Bay of Biscay, as well as for the peninsular part of the Ionian and Adriatic seas, these areas were omitted from the analyses. Thirty-eight large basins and 22 small basin groups were defined prior to the statistical analyses, resulting in a $60 \times 233$ presence–absence matrix.

**Table 1** Marine areas used in the present study, sources and referring code

| Marine region | Source | Code |
|---|---|---|
| Baltic Sea (continental coast) | ICES* | U1 |
| Baltic Sea (peninsular coast) | ICES* | U2 |
| Gulf of Finland | ICES | U3 |
| Gulf of Riga | ICES | U4 |
| Kattegat Sound and Belts | ICES | U5 |
| Skagerrak | ICES | U6 |
| North Sea (continental coast) | ICES* | U7 |
| North Sea (insular coast) | ICES* | U8 |
| Irish Sea and St George's Channel | ICES | U9 |
| Bristol Channel | ICES | U10 |
| English Channel (continental coast) | ICES* | U11 |
| English Channel (insular coast) | ICES* | U12 |
| Bay of Biscay (Spanish coast) | ICES* | U13 |
| Bay of Biscay (French coast) | ICES* | U14 |
| North-east Atlantic Ocean (< 40 W) | ICES | U15 |
| North-east Atlantic Ocean (> 40 W) | ICES* | U16 |
| Mediterranean Sea (Spanish coast) | FAO | U17 |
| Mediterranean Sea (French coast) | FAO | U18 |
| Tyrrhenian Sea | FAO | U19 |
| Adriatic Sea (continental coast) | FAO* | U20 |
| Adriatic Sea (peninsular coast) | FAO* | U21 |
| Ionian Sea (continental coast) | FAO* | U22 |
| Ionian Sea (peninsular coast) | FAO* | U23 |
| Aegean Sea | FAO | U24 |
| Black Sea | FAO | U25 |

*Adapted; see text for details.

*Relevance of the data set*

Before performing cluster analysis (see below) we first checked that there was meaningful information in the data, i.e. that species were not independently distributed into the basins (or groups of basins). To do so we randomly selected 116 species and performed a principal components analysis (PCA) on the resulting presence–absence matrix. Using the coordinates of the basins along the two first axes of the PCA, Euclidean distances were calculated. The same procedure was repeated using the 117 remaining species, leading to a second distance matrix between basins (or group of basins). If species were distributed independently into river basins there should be no correlation between the two distance matrices. Conversely a strong positive relationship was expected if a strong spatial or biogeographical structure was present in the data. The statistical significance of the correlation between the two distance matrices was tested using a randomized Mantel test (Dietz, 1983).

*Hierarchical clustering*

A PCA was first performed on the $60 \times 233$ presence–absence matrix, in order to reduce the data matrix to a lower number of dimensions while preserving the maximum amount of inertia. Then a Euclidian dissimilarity matrix was calculated based on PCA coordinates for each basin or group of basins, and submitted to a hierarchical cluster analysis using Ward's minimum variance method (Legendre & Legendre, 1998). The level of cut on the dendrogram was chosen to maximize the number of biogeographical regions while preserving the geographical coherence (i.e. geographically close basins or groups of basins classified closely in the dendrogram).

**Species–area relationships**

The drainage basin area (DBA) was collected for 369 rivers. To test for the effects of DBA and biogeographical region on the slope of the species–area relationships we built a model predicting the local species richness (LSR) as a function of DBA and a categorical variable 'biogeographical region' (BR). Because of the heteroscedasticity of the regression's residuals when using original data, a logarithmic transformation was applied to both DBA and LSR [log $(x + 1)$ for LSR due to null values].

Spatial autocorrelation may be present in these geographical data, leading to non-independence of the residuals and thus violating one of the assumptions of standard regression analysis (Cliff & Ord, 1981). Non-independence of the residuals may also result from similarity in species composition between river basins due to common history or common environmental features. As geographical distance and species similarity are generally inversely related at large spatial scales (Nekola & White, 1999), and because geographical distance between river basins it not easily defined, we only considered the non-independence between basins resulting from their common species composition. With this aim we performed a PCA on the presence–absence matrix and considered basin position on the plane defined by the two first axes, and then used this information to render residuals independent with regard to between-basin similarity in species composition. Regression coefficients were estimated using a generalized least squares procedure integrating the correlation between residuals (Venables & Ripley, 1997). An exponential correlation structure was chosen such that the correlation between two observations decreased exponentially with the distance between them. The distance used was the Euclidean distance between two basins in the plane defined by the first and second axis of the PCA. A quadratic trend surface mapping was also integrated into the model by using as explanatory variables basin coordinates along the first, PCA1, and second, PCA2, axis of the PCA, their squares, PCA1$^2$ and PCA2$^2$, and the interaction term PCA1 × PCA2. The combination of trend surface and spatially correlated residuals was recently shown to account for both broad-scale and fine-scale spatial structure (Lichstein *et al.*, 2002). Three nested models were fitted: (1) log (DBA) + quadratic trend, (2) log (DBA) + quadratic trend + BR, and (3) log (DBA) + quadratic trend + BR + log (DBA) × BR. A likelihood ratio test was used to assess if a more complex model significantly increased the fit with the data in comparison to the simple nested model. We tested for 'spatial' autocorrelation of the residuals from the retained model using Moran's *I* (Sokal & Oden, 1978). From inter-basin Euclidean distances in the PCA1–PCA2 plane, 20 distance classes were defined and the standardized *I* was

computed for each of them. Basins having exactly the same species composition (species list) have a null distance in the PCA plane. Unfortunately, null distances are not allowed to compute the exponential correlation matrix between river basins. Within groups of basins having the same species composition, only one basin was selected at random to be included into the regression analyses, leading to a sample size of 322 basins.

To graphically compare β-diversity (i.e. the dissimilarity in species composition between localities) among regions, 'Westoby's plots' (Westoby, 1993) were used. This method involves plotting species richness vs. area using values at the local and regional scale, the slope of the line connecting local to regional values being positively related to the β-diversity of the region (Westoby, 1993). Here localities were the hydrographical networks and regions the biogeographical regions. Local values were the mean of the log (LSR) and the mean of the log (DBA) for the networks within the region under consideration. Regional species richness (RSR) was the total number of species recorded within the region, and regional drainage basin area was the sum of the areas of the networks within the region.

Statistical analyses were performed using S+.

## RESULTS

### Fish community composition

The Cyprinidae family was strongly dominant with 121 species out of a total of 233, i.e. 51.9% of the total number of species (see Appendix S1 in Supplementary Material). Then, four families were represented by more than 10 species, i.e. 4.3% of the total number of species (Cobitidae, Gobiidae, Salmonidae and Percidae), and 10 families were represented by two to nine species, i.e. 0.9–3.9% of the total number of species (Acipenseridae, Clupeidae, Petromyzontidae, Balitoridae, Cottidae, Gasterosteidae, Cyprinodontidae, Atherinidae, Siluridae and Valenciidae). Finally, six families were represented by only one species, i.e. 0.4% of the total number of species: Anguillidae (*Anguilla anguilla*), Bleniidae (*Salaria fluviatilis*), Esocidae (*Esox lucius*), Gadidae (*Lota lota*), Osmeridae (*Osmerus eperlanus*) and Umbridae (*Umbra krameri*).

### Definition of biogeographical regions

The between-basin faunistic distances (Euclidean distances according to PCA axes, see Methods) obtained with a first set of 116 species were highly concordant with the ones obtained with the 117 other species ($r = 0.894$, $P < 0.001$, Mantel test with 1000 permutations), suggesting that there was a strong biogeographical structure in the data set. The cluster analysis allowed the definition of seven main biogeographical regions: Western Peri-Mediterranea, Central Peri-Mediterranea, Eastern Peri-Mediterranea, Ponto-Caspian Europe, Northern Europe, Central Europe and Western Europe (Figs 1 and 2). There was geographical coherence in the hierarchical clustering, i.e. geographically close catchments were classified close together in the dendrogram. The large Scandinavian catchments were an exception to this, being classified in Northern Europe.

### Patterns in species richness and endemism

Among the biogeographical regions, RSR was high for Central Peri-Mediterranea (110 species) and Ponto-Caspian Europe (98 species), while 64, 54, 52, 42 and 36 species were found in Eastern Peri-Mediterranea, Central Europe, Western Europe, Northern Europe and Western Peri-Mediterranea, respectively (Fig. 3 and Appendix S1 in Supplementary Material).

Many species (148 out of a total of 233, i.e. 63.5%) were endemic to only one biogeographical region. Four species, i.e. *A. anguilla*, *T. tinca*, *Gasterosteus aculeatus* and *S. trutta*, were found in all biogeographical regions. The percentage of endemics was high for Central Peri-Mediterranea (61.8%), intermediate for Western Peri-Mediterranea, Ponto-Caspian Europe and Eastern Peri-Mediterranea (44.4%, 36.7% and 31.2%, respectively), and low for Northern, Western and Central Europe (9.5%, 5.8% and 1.9%, respectively) (Fig. 3 and Appendix S1 in Supplementary Material). The endemism in central Peri-Mediterranea concerned many genera, among which were *Barbus* and *Chondrostoma* (seven species each), *Telestes* (six species) and Rutilus (five species). Considering Western Peri-Mediterranea, endemism mainly concerned three genera, i.e. *Chondrostoma* (five species), *Barbus* and *Squalius* (four species each), representing a total of 13 species out of a total of 15 endemics. For Ponto-Caspian Europe, endemism notably concerned *Neogobius*, *Romanogobio* and *Sabanejewia* (four, four and three species, respectively) but also typically migratory genera, i.e. *Acipenser* (four species) and *Alosa* (three species). Endemism in Eastern Peri-Mediterranea exhibited similarities with the Western Peri-Mediterranea, with *Chondrostoma* and *Cobitis* being well represented (three species each) (Appendix S1 in Supplementary Material).

### Species–area relationships

From the 369 available DBA, 80 corresponded to rivers situated in the Western and Central Peri-Mediterranea, 25 in the Eastern Peri-Mediterranea, seven in Ponto-Caspian Europe, 10 in Northern Europe, 119 in Western Europe and 128 in Central Europe.

Only the five biogeographical regions with enough available DBA data were considered for the present analyses, i.e. Western Peri-Mediterranea, Central Peri-Mediterranea, Eastern Peri-Mediterranea, Western Europe and Central Europe. For the likelihood ratio tests, the best models were those allowing for differences among regions in the species–area slopes: the increase in the goodness of fit due to the interaction term between BR and log (DBA) was highly significant (Table 2). The same conclusion was reached considering the Akaike information criterion (model 3 had the lowest value, Table 2). None of the standardized Moran's *I* computed on residuals of this model for 20 distance classes was higher in absolute value than 2 (Table 2). As a standardized Moran's *I* is considered significant at the 0.05 level if higher in absolute value than 1.96, this suggests that the lack of independence between residuals due to faunal similarity between basins was successfully removed by the model. For comparison, the equivalent non-spatial model, a simple least-square regression between log (LSR) and log (DBA), BR
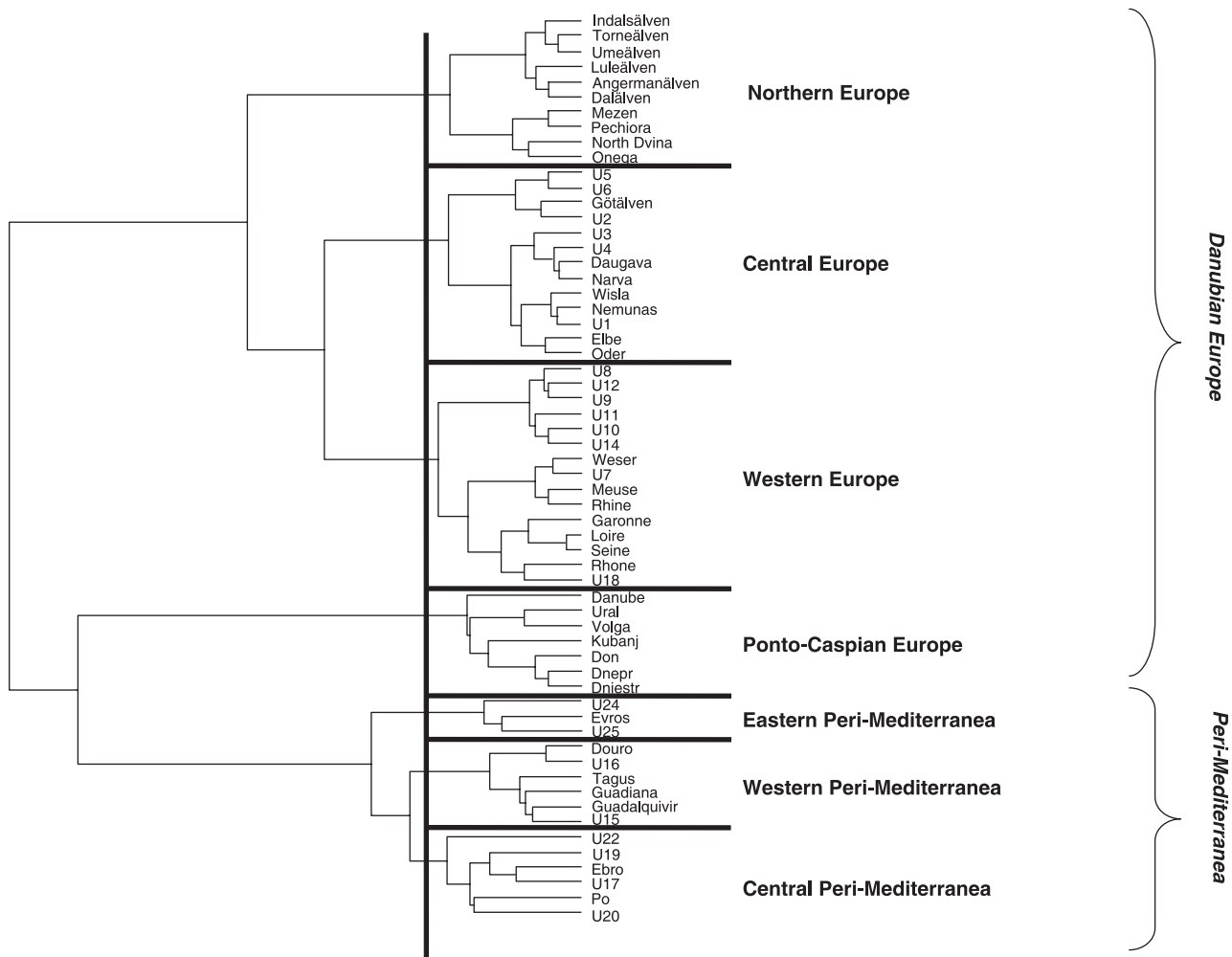
**Figure 1** Dendrogram from the hierarchical cluster analysis of the 38 large basins and 22 basin groups, showing the separation into seven biogeographical regions. The cutoff level was chosen in order to maximize the number of biogeographical regions while preserving geographical coherence between regions, and is illustrated by the thick vertical line.

**Table 2** Comparison of the generalized least squares models fitted to log (LSR)

| Model | AIC | Test | Lr | P value | I max | I min |
|---|---|---|---|---|---|---|
| Log (DBA) + trend | 170.2 | | | | | |
| Model 1 + BR | 139.3 | 1 vs. 2 | 38.86 | < 0.0001 | | |
| Model 2 + log (DBA) × BR | 90.3 | 2 vs. 3 | 57.06 | < 0.0001 | 1.79 | −1.62 |

LSR, local species richness; DBA, drainage basin area; BR, biogeographical region; trend, quadratic trend surface (according to coordinates along the two first axes of a PCA, see text); AIC, Akaike information criterion; Lr, likelihood ratio; I, standardized Moran's I (maximum and minimum values observed within 20 distance classes are provided).

and log (DBA) × BR, resulted in residuals that were highly autocorrelated with regard to faunal similarity between basins. For the 20 distance classes, the standardized Moran's I ranged from −5.88 to 3.47.

Eastern Peri-Mediterranea was characterized by high slope (0.43), Western Central Peri-Mediterranea by intermediate ones (0.39 and 0.40, respectively), and Central and Western Europe by low slopes (0.23 and 0.26, respectively) (Fig. 4).

According to Westoby's plots, the β-diversity was high for Central Peri-Mediterranea, intermediate for Western and Eastern Peri-Mediterranea and low for Central and Western Europe (Fig. 5).

## DISCUSSION

### Biogeographical regions

Seven biogeographical regions with homogeneous fish fauna (in terms of species presence–absence) were defined in the present work. To our knowledge, this is the first time a study has defined biogeographical regions on such a large spatial scale, using a statistical analysis.

The proposed classification has two main limits. First, some large Swedish basins were classified in Northern Europe, together with the Mezen, North Dvina, Onega and Pechiora basins. This
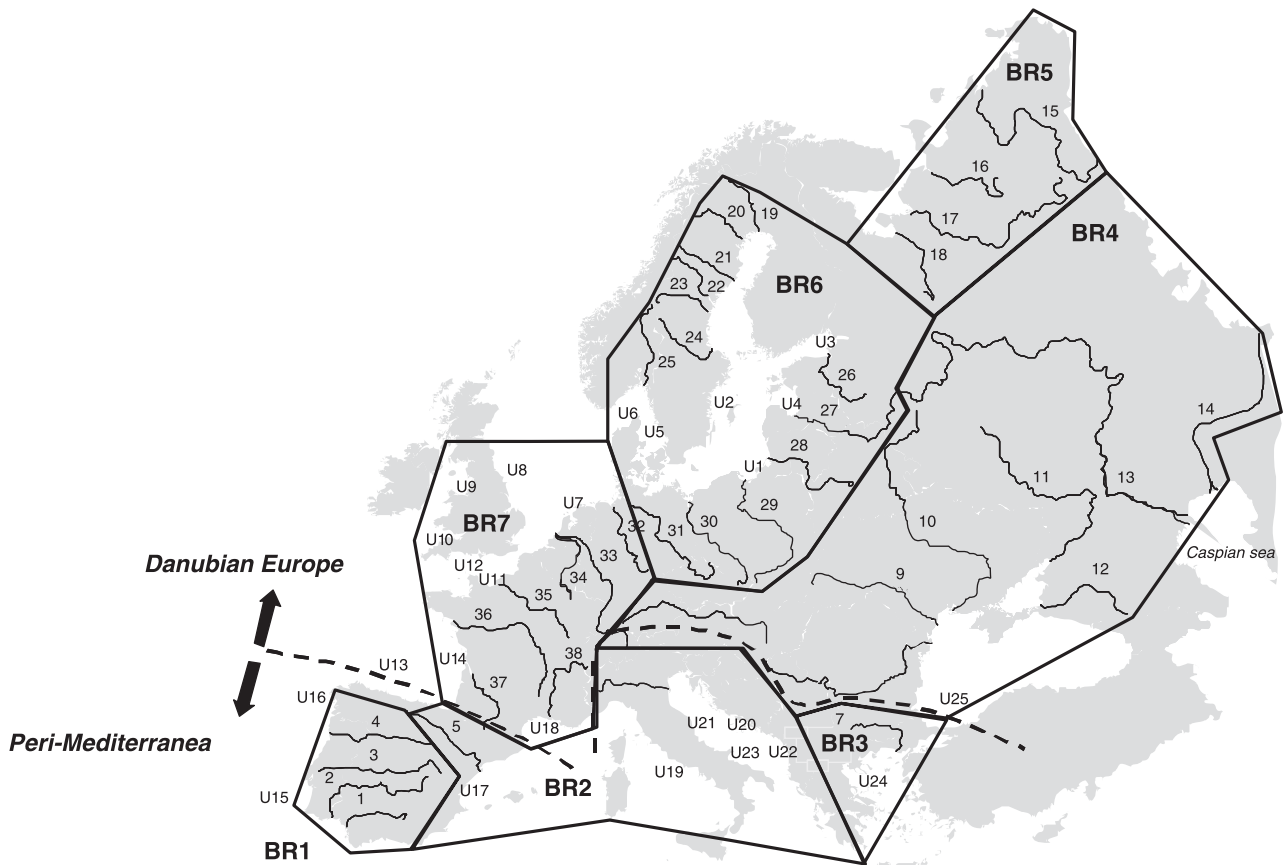
**Figure 2** Map showing the 38 large rivers and 25 fishing areas (see text for details): 1, Guadalqivir; 2, Guadiana; 3, Tagus; 4, Douro; 5, Ebro; 6, Pô; 7, Evros; 8, Danube; 9, Dniestr; 10, Dniepr; 11, Don; 12, Kubanj; 13, Volga; 14, Ural 15, Pechiora; 16, Mezen; 17, Dvina; 18, Onega; 19, Torneälven; 20, Luleälven; 21, Umeälven; 22, Angermanälven; 23, Indalsälven; 24, Dalälven; 25, Gota älven; 26, Narva; 27, Daugava; 28, Nemunas; 29, Wisla; 30, Oder; 31, Elbe; 32, Weser; 33, Rhine; 34, Meuse; 35, Seine; 36, Loire; 37, Garonne; 38, Rhône. The boundaries of the seven biogeographical regions are illustrated, with: BR1, Western Peri-Mediterranea; BR2, Central Peri-Mediterranea; BR3, Eastern Peri-Mediterranea; BR4, Ponto-Caspian Europe; BR5, Northern Europe; BR6, Central Europe; BR7, Western Europe. The dashed line shows the separation between Danubian Europe and Peri-Mediterranea.

geographical incoherence was notably related to species richness, the richness of large Swedish catchments (except Götaalven) being small compared with other large catchments from Central Europe (e.g. the Elbe, Nemunas, Oder and Wisla), but close to the species richness of large catchments from Northern Europe. Thus, the methodology used in the present work (i.e. considering only large basins > 25,000 km$^2$ independently and pooling smaller ones) did not perfectly eliminate the potential bias due to basin size effect. However, given that geographical coherence was respected for most of the 38 large basins and 22 basin groups, our results appear reliable even if probably unprovable. Secondly, it is important to keep in mind that the spatial scale considered did not exactly match a biogeographical unit, since faunal similarities exist between Southern European basins (i.e. Spain and Greece) and North African and Middle Eastern basins (Banarescu, 1990, 1992; Banarescu & Coad, 1991). Adding data from these regions and thus considering the whole biogeographical unit would probably lead to the enlargement of some of the biogeographical regions defined herein, or to defining new ones.

Compared with the heuristic checklist of the freshwater fish of Europe (Kottelat, 1997), and except for the genera *Coregonus* and *Salmo* because of the taxonomical complexity of these two groups, 219 species were used in the present work compared with 262 in Kottelat (1997). The comparison of these two works is sometimes difficult because of taxonomical uncertainties, but most species used in this work can also be found in Kottelat (1997). Therefore, the species list used in the present work, although incomplete, can be considered as representative of the whole European freshwater fish fauna.

## Patterns in species richness and endemism

On the European scale, the Cyprinidae family was the most well represented, with more than 50% of the total number of species. Cyprinidae is the largest family of freshwater fish and is widely distributed on Earth, with representatives in Eurasia, North America and Africa (Banarescu, 1990; Winfield & Nelson, 1991). Traditionally, East and especially South-East Asia (the part of Asia situated to the south-east of the Tibetan Plateau) are often
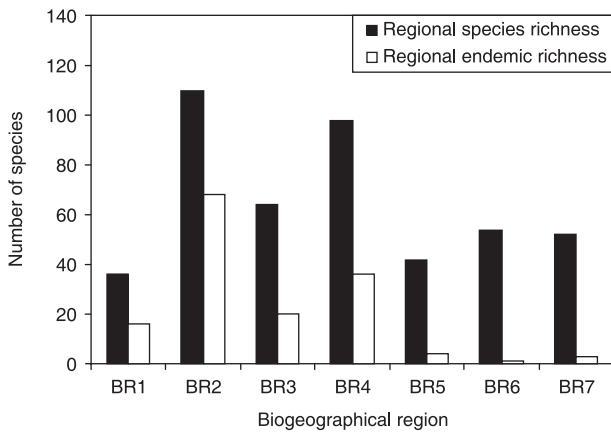
Figure 3 Regional species richness (RSR) and regional endemic richness (RER) for the seven biogeographical regions: BR1, Western Peri-Mediterranea; BR2, Central Peri-Mediterranea; BR3, Eastern Peri-Mediterranea; BR4, Ponto-Caspian Europe; BR5, Northern Europe; BR6, Central Europe; BR7, Western Europe.
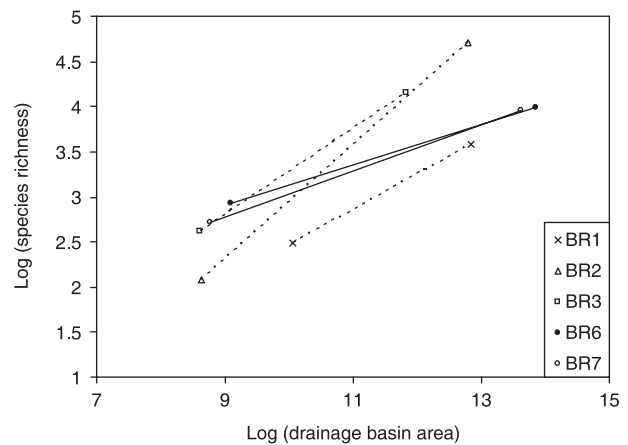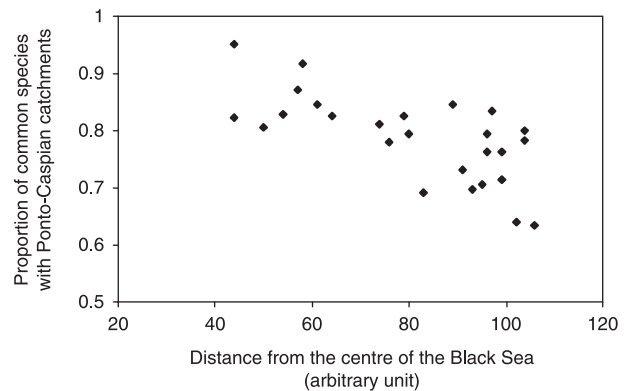


Figure 5 Westoby's plots for Western Peri-Mediterranea (BR1), Central Peri-Mediterranea (BR2), Eastern Peri-Mediterranea (BR3), Central Europe (BR6) and Western Europe (BR7). Dashed lines: Peri-Mediterranea.
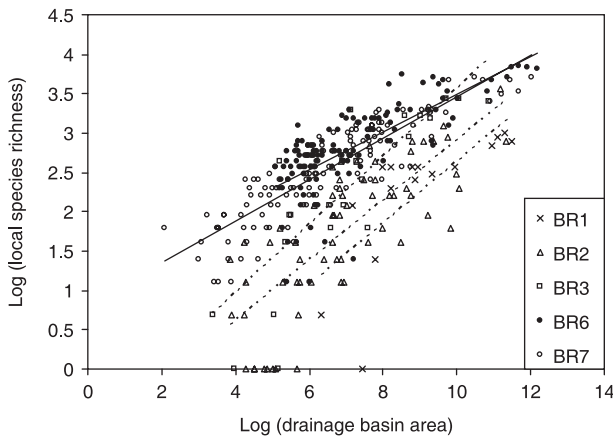


Figure 4 Species–area relationships for Western Peri-Mediterranea (BR1), Central Peri-Mediterranea (BR2), Eastern Peri-Mediterranea (BR3), Central Europe (BR6) and Western Europe (BR7). Dashed lines: Peri-Mediterranea.



Figure 6 Abscissa: approximate distance (arbitrary units) between large catchments from Northern, Central and Western Europe and the centre of the Black Sea. Ordinates: proportion of common species between large catchments from Northern, Central and Western Europe and large catchments from Ponto-Caspian Europe (Danube, Don, Dnepr, Dniestr and Kubanj). Danube, Don, Dnepr, Dniestr and Kubanj are grouped together because the Black Sea has been likely to allow fish dispersal among Ponto-Caspian catchments due to its low salinity. Ural and Volga are not included in Ponto-Caspian catchments because of the common past history between the Black and Caspian seas before the LGM (common species could thus not necessarily indicate recolonization).

viewed as the diversification and dispersal centre for Cypriniformes, while South-West Asia, Siberia, Europe and North America were colonized afterwards (see Banarescu, 1990, 1992; Bianco, 1990; Winfield & Nelson, 1991; for details).

For Europe, the main event that affected Europe during the Quaternary era was a succession of numerous cold and warm periods (Bennett, 1997; Williams et al., 1998; Hewitt, 1999, 2000). The main surviving species corresponded to those present in the Danube Basin before the glaciations, whose fauna derived from the freshwater or oligosaline phase of the Paratethys during the Miocene (the 'Lago Mare' phase), allowing the first dispersion of Cyprinidae and other primary fish from Central Asia to eastern and Central Europe (Bianco, 1990; Banarescu, 1990, 1992; Williams et al., 1998). Therefore, it is usually supposed that the uniform cyprinid fauna of the upper part of Europe (north of

mountain chains), depleted during the glacial periods, was recolonized with Danubian fish during the interglacial and post-glacial periods (Banarescu, 1990, 1992). This explains why, among the seven main regions defined here, Central, Western and Northern Europe were characterized by low species richness, while Ponto-Caspian Europe was characterized by high species richness. This hypothesis is also well supported by Fig. 6, which shows a decrease in the proportion of common species between large catchments from Central, Western and Northern Europe

and large catchments from Ponto-Caspian Europe as the distance from Ponto-Caspian Europe increases (logistic regression, $P < 0.001$). The high species richness of Ponto-Caspian Europe might also be directly related to the oligosaline or freshwater condition of Paratethys, which allowed many marine species, notably Gobiidae, to adapt to freshwater.

Regarding endemism, the high level found in Ponto-Caspian Europe can also be related to the role played by the Danube river basin during the last glaciations. A different history may explain the high endemism observed in Peri-Mediterranean areas (more than 60% of the species in Central Peri-Mediterranea more than 30% in the Western Peri-Mediterranea and Eastern Peri-Mediterranea). Three hypotheses have been previously invoked to explain the origin and diversity of endemic freshwater fishes in the Mediterranean peninsulas (see Sanjur *et al.*, 2003 for details). The first postulated that the spread of cyprinids to the southern part of Europe and North Africa across the Gibraltar strait was possible until the Pliocene (i.e. 2–5 Ma) and that subsequent isolation of the Iberian Peninsula and southern Greece would have been responsible for their rich endemic fauna (Almaça, 1976; Banarescu, 1989, 1992). The second suggested that dispersions occurred through intercontinental land bridges during the formation of the actual North African coast in the Early Pliocene (Doadrio, 1990). The last referred to the Messinian salinity crisis (about 5 Ma), during which the Mediterranean Basin almost dried up and was refilled with freshwater from 'Lago Mare' Paratethys (Hsü *et al.*, 1977), and would have allowed the dispersal of cyprinids throughout the circum-Mediterranean regions (Bianco, 1990). These three hypotheses are not mutually exclusive, and probably all contribute to the high level of endemism found in Peri-Mediterranean areas. However, the 'Lago Mare' hypothesis was supported by the absence of fossil remains of fish related to the fauna of Danubian Europe until the Messinian (Greenwood, 1974; Sorbini & Tirapelle Rancan, 1979; Bianco, 1998), and has been tested recently by several studies on molecular genetics, either on Danubian and Peri-Mediterranean taxa of Cyprinidae in general (Gilles *et al.*, 1998), or in a single genus such as *Chondrostoma* (Durand *et al.* 2003), *Scardinius* and *Telestes* (Ketmaier *et al.* 2004).

In contrast to Peri-Mediterranean regions, where endemism was high, endemism was very low in Northern, Central and Eastern Europe. This was probably due to vicariance/dispersal events which occurred repeatedly in western Palaearctic areas during glaciations, and concords with the previous work by Oberdorff *et al.* (1999), which showed that partially or fully glaciated European areas during the Pleistocene shelter few endemic species today. During interglacial periods, melt-water drainage systems were in operation connecting most Central European and Siberian river basins (Thunnel & Williams, 1983), and pro-glacial meltwater lakes have been shown previously to offer large dispersal opportunities (Hocutt & Wiley, 1986). Several primary species probably used these systems to disperse or recolonize rivers at glacial–interglacial intervals. An indicator in favour of this hypothesis is that the chub, *Squalius cephalus*, and the common barbel, *Barbus barbus*, widespread in these three regions, were shown to exhibit little genetic variation among populations very far from each other (Durand

*et al.*, 1999; Kotlik & Berrebi, 2001). Similar findings were highlighted for North American freshwater fish (Bernatchez & Wilson, 1998).

## Species–area relationships

Two main sampling schemes have been used previously to generate species–area relationships: nested sampling designs for mainland areas and independent sampling designs for archipelagos, with islands as samples (Rosenzweig, 1995). Considering riverine fish, independent sampling designs have been favoured, following two different approaches. The first one considered that a stream could be viewed as an archipelago *per se*, with morphological stream units, i.e. riffles and pools representing islands (Angermeier & Schlosser, 1989). The second approach, developed here, considers the fluvial networks of drainage basins as biogeographical islands containing specific pools of species (Livingstone *et al.*, 1982; Hugueny, 1989a,b).

Three main hypotheses are usually postulated to explain species–area relationships: the habitat diversity hypothesis (Williams, 1964), the area *per se* hypothesis (Simberloff, 1976) and the passive sampling hypothesis (Connor & McCoy, 1979). It is noteworthy that none of these traditional explanations of the species–area relationship takes the role of speciation, extinction and immigration/emigration rates into account, and are therefore necessarily incomplete (MacArthur & Wilson, 1963; Rosenzweig, 1995; Connor & McCoy, 2001).

Based on the present study, several observations can be made with regard to species–area relationships. First, two groups of plots were obtained: a Danubian one (Central and Western Europe), with low slope values, and a Peri-Mediterranean one, with intermediate or high slopes. Second, the species–area relationships observed for Danubian Europe were above the ones of Peri-Mediterranean areas. Similarly, two groups of areas were identified by Westoby's plot: a Danubian one, with low slopes, and a Peri-Mediterranean one, with intermediate or high slopes.

Excepting the Messinian salinity crisis and Pleistocene glaciations, during which dispersal and faunal exchanges were probably of exceptional importance, river basins can be considered as non-equilibrated islands in which population extinctions are not fully balanced by immigration from surrounding basins. Empirical evidence (Brown, 1971) and theoretical work (Wright, 1981) have shown that for such systems the slope $z$ of the species–area relationship in a log–log representation increases as time elapses from initial isolation, because extinction is more likely in small islands (Fig. 7a). Similar results apply when comparing species–area slopes of recent vs. old lineages within the same archipelago (Ricklefs & Bermingham, 2004). Given this framework, it is expected that $z$ should be higher in the Peri-Mediterranean regions than in the northern ones because the last major faunal exchange is older for the former (5 Ma) than for the latter (10,000 years ago). Our results were concordant with this hypothesis: the lowest slope was observed for Central and Western Europe, while Peri-Mediterranean regions exhibited the highest slopes.

On the other hand, if a river basin is isolated from its neighbours for a long time, allopatric speciation is likely to occur.
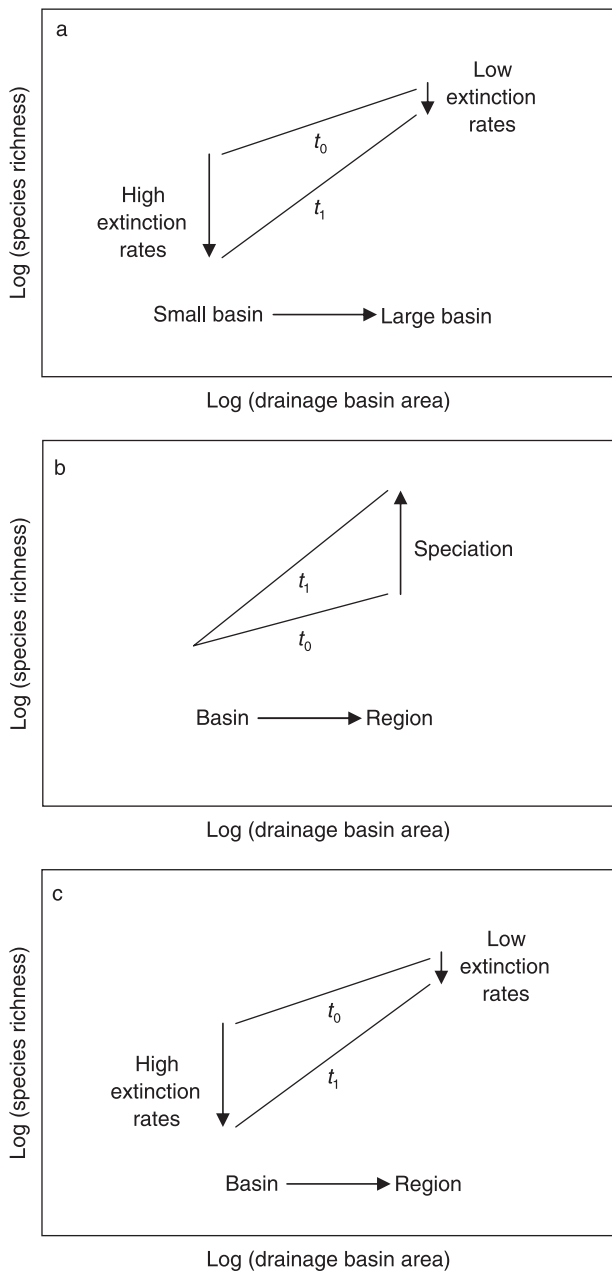
Figure 7 (a) Theoretical effect of extinction on the species–area relationships if river basins are completely isolated between $t_0$ (last faunal exchange between basins) and $t_1$ (present) and if the probability of extinction is higher in small than in large basins or within a basin compared to within all the basins of the region. (b) Theoretical effect of allopatric speciation on Westoby's plot if river basins are completely isolated between $t_0$ (last faunal exchange between basins) and $t_1$ (present). (c) Theoretical effect of extinction on Westoby's plots if river basins are completely isolated between $t_0$ (last faunal exchange between basins) and $t_1$ (present) and if the probability of local extinction (within one river) is higher than the probability of regional extinction (within all the rivers of the region).

Allopatric speciation is defined here as the speciation resulting from the isolation of populations into different basins. A consequence of allopatric speciation can be an increase in regional species richness without changes in the number of species per basin, in the absence of extinction phenomena (Stephens & Wiens, 2003) (e.g. a species present in 10 different basins at $t_0$ which gives 10 different species at $t_1$). Consequently, as a Westoby's plot links local species richness (per hydrographical network) to the regional one, the slope of the species–area relationship should increase (Fig. 7b). Considering extinction, extinction in one river is more likely than simultaneous extinction in all the rivers within a region. Thus, extinctions should also increase the slope of Westoby's plots (Fig. 7c). Combining the supposed effects of extinction and speciation we expect that the longer the rivers of a region have been isolated, the higher the slope of Westoby's plots should be. This expectation was fulfilled, as shown by Fig. 4: slopes were steeper for Peri-Mediterranean regions than for Danubian Europe (Central and Western Europe). With regard to the other regions, Central and Eastern Peri-Mediterranea were characterized by high regional species richness. This suggests that a high number of *in situ* speciations occurred within these regions and this is consistent with their high number of regional endemics.

For similarly sized basins, Peri-Mediterranean rivers harbour fewer species than northern ones, as expected if they had been subjected to extinction processes for a longer time. However, this historical explanation should be considered with caution because the low diversity observed in Peri-Mediterranean rivers may be due to their present harsh hydrological conditions (see Gasith & Resh, 1999 for details).

It is important to keep in mind that the present study did not address a species–area relationship for Ponto-Caspian Europe because of the absence of data for small catchments. This would have been of major interest regarding the major role played by this area during the LGM, and because of the high species richness and endemism observed there.

A recent paper by Storch *et al.* (2005) showed that there was a negative interaction between energy availability and area in their effect on species richness: the slope of the species–area relationship was lower in areas with higher levels of available energy. This phenomenon could have played a role here, and differences in species richness may reflect differences in climate. This was partly taken into account, since past climatic events such as glaciations were believed to be responsible for some of the patterns observed. However, further studies using additional data would be necessary to address the issue.

## CONCLUSION

The biogeographical regions defined in this study were congruent with the known biogeographical history of European freshwater fish, notably the Messinian salinity crisis and Pleistocene glaciations. As a consequence, they should now be used to address important ecological questions, since they add a relevant spatial scale to investigation (between continents and hydrographical networks), as well as to management expectations, as recommended by Hughes *et al.* (1994) and Matthews (1998). Finally, we suggest that Peri-Mediterranean areas and Ponto-Caspian Europe should be considered as 'biodiversity hotspots' for European freshwater fish.

## ACKNOWLEDGEMENTS

## REFERENCES

Almaça, E.C. (1976) La speciation chez les Cyprinidae de la Péninsule ibérique. *Revue des Travaux de l'Institut des Pêches Maritimes*, **40**, 399–411.

Angermeier, P.L. & Schlosser, I.J. (1989) Species-area relationships for stream fishes. *Ecology*, **70**, 1450–1462.

Banarescu, P. (1989) Zoogeography and history of the freshwater fish fauna of Europe. *The freshwater fishes of Europe, Volume 1. General introduction to fishes, Acipenseriformes* (ed. by J. Holčik), pp. 88–107. AULA-Verlag, Wiesbaden.

Banarescu, P. (1990) *Zoogeography of fresh waters. Volume 1. General distribution and dispersal of freshwater animals.* AULA-Verlag, Wiesbaden.

Banarescu, P. (1992) *Zoogeography of fresh waters. Volume 2. Distribution and dispersal of freshwater animals in North America and Eurasia.* AULA-Verlag, Wiesbaden.

Banarescu, P. & Coad, B.W. (1991) Cyprinids of Eurasia. *Cyprinid fishes. Systematics, biology, and exploitation* (ed. by I.J. Winfield and J.S. Nelson), pp. 127–155. Chapman & Hall, London.

Bennett, K.D. (1997) *Evolution and ecology: the pace of life.* Cambridge University Press, Cambridge.

Bernatchez, L. & Wilson, C.C. (1998) Comparative phylogeography of Nearctic and Palearctic fishes. *Molecular Ecology*, **7**, 431–452.

Bianco, P.G. (1990) Potential role of the paleohistory of the Mediterranean and Paratethys basins on the early dispersal of Euro-Mediterranean freshwater fishes. *Ichthyological Exploration of Freshwaters*, **1**, 167–184.

Bianco, P.G. (1995) Factors affecting the distribution of freshwater fishes especially in Italy. *Cybium*, **19**, 241–259.

Bianco, P.G. (1998) Diversity of Barbinae fishes in southern Europe with description of a new genus and a new species. *Italian Journal of Zoology*, **65** (Suppl.), 125–136.

Bilton, D.T., Mirol, P.M., Mascheretti, S., Fredga, K., Zima, J. & Searle, J.B. (1998) Mediterranean Europe as an area of endemism for small mammals rather than a source for northwards postglacial colonization. *Proceedings of the Royal Society of London Series B*, **265**, 1219–1226.

Blondel, J. & Vigne, J.-D. (1993) Space, time and man as determinants of diversity of birds and mammals in the Mediterranean Region. *Historical and geographical determinants of community diversity* (ed. by R.E. Ricklefs and D. Schluter), pp. 135–146. Chicago University Press, Chicago.

Brown, J.H. (1971) Mammals on mountaintops: nonequilibrium insular biogeography. *The American Naturalist*, **105**, 467–478.

Cliff, A.D. & Ord, J.K. (1981) *Spatial processes: models and applications.* Pion Ltd, London.

Conner, J.V. & Suttkus, R.D. (1986) Zoogeography of freshwater fishes of the western Gulf Slope of North America. *The zoogeography of North American freshwater fishes* (ed. by C.H. Hocutt and E.O. Wiley), pp. 413–456. John Wiley & Sons, New York.

Connor, E.F. & McCoy, E.D. (1979) Species-area relationships. *Encyclopedia of Biodiversity*, **5**, 397–411.

Connor, E.F. & McCoy, E.D. (2001) The statistics and biology of the species-area relationship. *The American Naturalist*, **113**, 791–833.

Dietz, J.E. (1983) Permutation tests for association between two distances matrices. *Systematic Zoology*, **32**, 21–26.

Doadrio, I. (1990) Phylogenetic relationships and classification of west Palearctic species of the genus Barbus (Osteichthyes: Cyprinidae). *Aquatic Living Resources*, **3**, 265–282.

Durand, J.-D., Bianco, P.G., Laroche, J. & Gilles, A. (2003) Insight into the origin of endemic Mediterranean ichthyofauna. Phylogeography of *Chondrostoma* genus (Teleostean, Cyprinidae). *Journal of Heredity*, **94**, 315–328.

Durand, J.-D., Persat, H. & Bouvet, Y. (1999) Phylogeography and postglacial dispersion of the chub (*Leuciscus cephalus*) in Europe. *Molecular Ecology*, **8**, 989–997.

Gasith, A. & Resh, V.H. (1999) Streams in Mediterranean climate regions: abiotic influences and biotic responses to predictable seasonal events. *Annual Review of Ecology and Systematics*, **30**, 51–81.

Gilbert, C.R. (1980) Zoogeographic factors in relation to biological monitoring of fish. *Biological monitoring of fish* (ed. by C.H. Hocutt and J.R. Stauffer Jr), pp. 309–339. Lexington Books, Lexington.

Gilles, A., Lacointre, G., Faure, E., Chappaz, R. & Brun, G. (1998) Mitochondrial phylogeny of the European cyprinids: implications for their systematics, reticulate evolution, and colonization time. *Molecular Phylogenetics and Evolution*, **10**, 132–143.

Greenwood, P.H. (1974) Review of Cenozoic freshwater fish fauna in Africa. *Annals of the Geological Survey of Egypt*, **4**, 211–232.

Hewitt, G.M. (1999) Post-glacial recolonization of European biota. *Biological Journal of the Linnean Society*, **68**, 87–112.

Hewitt, G.M. (2000) The genetic legacy of the Quaternary ice ages. *Nature*, **405**, 907–913.

Hocutt, C.H. & Wiley, E.O. (1986) *The zoogeography of North American freshwater fishes.* John Wiley & Sons, New York.

Hsü, K., Montadert, L., Bernuilli, D., Cita, M.B., Erickson, A., Garrison, R.E., Kidd, R.B., Mèlierés, F., Müller, C. & Wright, R. (1977) History of the Mediterranean salinity crisis. *Nature*, **267**, 399–403.

Hughes, R.M., Heiskary, S.A., Matthews, W.J. & Yoder, C.O. (1994) Use of ecoregions in biological monitoring. *Biological monitoring of aquatic systems* (ed. by S.L. Loeb and A. Spacie), pp. 125–151. Lewis Publishers, Boca Raton.

Hugueny, B. (1989a) *Biogéographie et structure des peuplements de poisons d'eau douce de l'Afrique de l'Ouest: approches quantitatives.* Thèse de Doctorat, Université Paris 7, Paris.

Hugueny, B. (1989b) West African rivers as biogeographic islands: species richness of fish communities. *Oecologia*, **79**, 236–243.

Ketmaier, V., Bianco, P.G., Cobolli, M., Krivokapoic, M., Caniglia, R. & De Matthaeis, E. (2004) Molecular phylogeny of two lineages of Leuciscinae cyprinids (*Telestes* and *Scardinius*) from the peri-Mediterranean area based on cytochrome b data. *Molecular Phylogenetics and Evolution*, **32**, 1061–1071.

Kotlik, P. & Berrebi, P. (2001) Phylogeography of the barbell (*Barbus barbus*) assessed by mitochondrial DNA variation. *Molecular Ecology*, **10**, 2177–2185.

Kottelat, M. (1997) European freshwater fishes: an heuristic checklist of the freshwater fishes of Europe (exclusive of former USSR), with an introduction for non-systematists and comments on nomenclature and conservation. *Biologia (Bratislava)*, **52** (Suppl. 5), 1–271.

Legendre, P. & Legendre, L. (1998) *Numerical ecology*, 2nd English edn. Elsevier Science BV, Amsterdam.

Lichstein, J.W., Simons, T.R., Shriner, S.A. & Franzreb, K.E. (2002) Spatial autocorrelation and autoregressive models in ecology. *Ecological Monographs*, **72**, 445–463.

Livingstone, D.A., Rowland, M. & Bailey, P.E. (1982) On the size of African riverine fish faunas. *American Zoologist*, **22**, 361–369.

MacArthur, R.H. & Wilson, E.O. (1963) An equilibrium theory of insular zoogeography. *Evolution*, **17**, 373–387.

Matthews, W.J. (1998) *Patterns in freshwater fish ecology*. Chapman & Hall, New York.

Moyle, P.B. & Herbold, B. (1987) Life-history patterns and community structure in stream fishes of western North America: Comparisons with eastern North America and Europe. *Community and evolutionary ecology of North American stream fishes* (ed. by W.J. Matthews and D.C. Heins), pp. 25–32. University of Oklahoma Press, Oklahoma.

Myers, G.S. (1951) Freshwater fishes and East Indian zoogeography. *Stanford Ichthyological Bulletin*, **4**, 11–21.

Nekola, J.C. & White, P.S. (1999) The distance decay of similarity in biogeography and ecology. *Journal of Biogeography*, **26**, 867–878.

Oberdorff, T., Lek, S. & Guégan, J.-F. (1999) Patterns of endemism in riverine fish of the Northern Hemisphere. *Ecology Letters*, **2**, 75–81.

Ricklefs, R.E. & Bermingham, E. (2004) History and the species-area relationship. *The American Naturalist*, **163**, 227–239.

Rosenzweig, M.L. (1995) *Species diversity in space and time*. Cambridge University Press, Cambridge.

Sanjur, O.I., Carmona, J.A. & Doadrio, I. (2003) Evolutionary and biogeographical patterns within Iberian populations of the genus *Squalius* inferred from molecular data. *Molecular Phylogenetics and Evolution*, **29**, 20–30.

Simberloff, D. (1976) Experimental zoogeography of islands: effects of island size. *Ecology*, **57**, 629–648.

Slechtova, V., Bohlen, J., Freyhof, J., Persat, H. & Del Mastro, G. (2004) The alps as barrier to dispersal of cold-adapted freshwater fishes? Phylogeographic history and taxonomic status of the bullhead in the Adriatic freshwater drainage. *Molecular Phylogenetics and Evolution*, **33**, 225–239.

Soffer, O. (1990) The Russian plain at the last glacial maximum. *The world at 18,000 BP. 1. High latitudes* (ed. by O. Soffer and C. Gamble), pp. 228–252. Unwin-Hyman, London.

Sokal, R.R. & Oden, N.L. (1978) Spatial autocorrelation in biology. 1. Methodology. *Biological Journal of the Linnean Society*, **10**, 199–228.

Sorbini, L. & Tirapelle Rancan, R. (1979) Messinian fossil fish of the Mediterranean. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **29**, 143–154.

Stephens, P.R. & Wiens, J.J. (2003) Species richness in emydid turtles. *The American Naturalist*, **161**, 112–128.

Storch, D., Evans, K.L. & Gaston, K.J. (2005) The species-area-energy relationship. *Ecology Letters*, **8**, 487–492.

Thunnel, R.C. & Williams, D.F. (1983) Paleotemperature and paleosalinity history of the eastern Mediterranean during the Late Quaternary. *Palaeogeography, Palaeoclimatology, Palaeoecology*, **44**, 23–29.

Venables, W.N. & Ripley, B.D. (1997) *Modern applied statistics with S-PLUS*, 2nd edn. Springer, New York.

Westoby, M. (1993) Biodiversity in Australia compared to other continents. *Historical and geographic determinants of community diversity* (ed. by R.E. Ricklefs and D. Schluter), pp. 170–177. University of Chicago Press, Chicago.

Williams, C.B. (1964) *Patterns in the balance of nature*. Academic Press, London.

Williams, M.A.J., Dunkerley, D.L., De Deckker, P., Kershaw, A.P. & Chappell, J. (1998) *Quaternary environments*, 2nd edn. Edward Arnold, London.

Winfield, I.J. & Nelson, J.S. (1991) *Cyprinid fishes. Systematics, biology and exploitation*. Chapman & Hall, London.

Wright, S.J. (1981) Intra-archipelago vertebrate distributions: the slope of the species-area relation. *The American Naturalist*, **118**, 726–748.

---

**BIOSKETCH**

**Yorick Reyjol** is a post-doctoral research associate at the University of Lyon (France). His research topics include fish community ecology, analyses of spatial and temporal patterns in rivers and lakes, effects of global change and modelling.

---

Editor: Pablo Marquet

**SUPPLEMENTARY MATERIAL**

The following material is available online at www.blackwell-synergy.com/loi/geb

**Appendix S1** Native fish faunistic list.