INTRODUCTION

Preserving species diversity is of major concern in ecology and conservation biology because biodiversity has both direct and indirect benefits upon which human welfare depends (e.g. Fuller et al. 2007). Most current management programs for biodiversity are largely pattern-based and focus on hotspots (Roberts et al. 2002). However, in a world increasingly transformed by human activities, process-based approaches are needed to anticipate biodiversity trends following global change. Unfortunately, a consensus about the main underlying mechanisms that shape the geographical distribution of species diversity is still elusive. This is particularly true for marine environments that suffer from a lack of emergent macroscale ecological knowledge and data (Spalding et al. 2007, Halpern & Floeter 2008), but see Connolly et al. (2003) and Gray (2002) for large-scale studies of species richness patterns for reef fishes, corals and benthos.

The foundations underlying fish diversity distribution in the Mediterranean Sea, which is under increasing human pressure, deserve particular attention due to the importance of conserving this hotspot of bio-

ABSTRACT: Many ecological drivers have been proposed to explain the spatial distribution of species richness (productivity, energy, climate), but historical and paleogeographical contingencies may also be important. All these hypotheses make the implicit assumption that species richness would be similar everywhere in absence of any direct environmental variations; this is known as the mid-domain effect (MDE). The Mediterranean Sea is an ‘archetypal situation’ where we would expect this effect, so we used the MDE as a null model for fish richness and analyzed the deviations. Temperature, primary production and distance from both the Gibraltar Strait and Suez Canal have been included in our models as predictors. To obtain our objective we (1) built a database containing the geographic distributions of 619 Mediterranean fishes, (2) implemented 1- and 2-dimensional MDE models, and (3) carried out spatial autoregressive models on MDE deviations. It appeared that in 1 dimension, the only true MDE was displayed longitudinally by the large-ranged endemic species. In 2 dimensions, the MDE accounted for 21% of variation in fish richness distribution on the continental shelf. However, we showed that neglecting the MDE when relating fish richness to environmental predictors is misleading. In fact, minimum temperature and primary production negatively influenced fish species richness per se, but had a positive effect on deviations from MDE predictions. Finally, our study suggested that by analyzing the deviations from MDE predictions and taking into account spatial autocorrelation we developed a more rigorous assessment of the factors influencing diversity patterns.

KEY WORDS: Null model · Autoregressive models · Species pattern · Geometric constraints · Primary production · Temperature · Geographical Information System · Mediterranean Sea
diversity and endemism (4 to 18% of all known marine species and 8.8% of endemism according to Quignard & Tomasini 2000). The main routes contributing to fish species richness in the Mediterranean Sea are the Strait of Gibraltar and the Suez Canal. The former has allowed Atlantic Ocean fauna to enter the Mediterranean Sea for 5 million years. The latter has allowed Lessepsian species to enter since 1869 (see Ben Rais Lasram & Mouillot 2008 and Ben Rais Lasram et al. 2008 for meta-analyses). As with most faunistic groups in the Mediterranean Sea (e.g. Danovaro et al. 2008), fish richness pattern is assumed to follow an eastward decline (Quignard & Tomasini 2000), but this trend was not confirmed by a recent study based on MEDITS surveys (surveys carried out on the continental shelf of the whole northern Mediterranean Sea; Gaertner et al. 2007). Despite these investigations, spatial patterns of fish diversity at the whole Mediterranean scale and their underlying mechanisms have been largely overlooked and no macroecological theory has been tested on the entire fish fauna.

Many ecological drivers have been proposed to explain the spatial distribution of species richness such as productivity and energy (Currie 1991), climate (Kerr et al. 1998) and habitat heterogeneity (Kerr et al. 2001) (see a list of theories in Rahbek & Graves 2001), but with a lack of consensus about the relative contribution of each of these factors. For instance, one of the most popular hypotheses is the general richness–climate relationship (Currie et al. 2004) that may influence fish richness by the mean of at least 2 drivers. Primary production ultimately limits the energy flow through marine food webs and, by cascade, the population growth of species and then species richness (Srivastava & Lawton 1998) (‘energy-richness hypothesis’). Water temperature is related to the ambient energy required by fishes to maintain metabolism (Ibarz et al. 2007) or avoid size limitation (Sonin et al. 2007) (‘physiological tolerance hypothesis’). However, environmental factors are not the only potential drivers of species richness patterns: historical, paleogeographical and even pure geometrical contingencies may also represent alternative explanations (Heads 2003, Storch et al. 2006).

All the hypotheses involving environmental variables as predictors of species richness make the implicit assumption that species richness would be similar everywhere in absence of direct environmental variations. Colwell & Hurtt (1994) challenged this assumption by proposing that the geometry of a biogeographic domain is of primary importance in explaining species richness patterns and that environmental variables are only of secondary importance. Indeed, one of the most intriguing and controversial theories that have been proposed to explain the spatial distribution of species richness is the mid-domain effect (MDE). Contrary to the classical ecological theories, the MDE suggests that geometric constraints on species distributions may create spatial patterns in species richness in the absence of any direct environmental gradient (Colwell & Hurtt 1994, Willig & Lyons 1998, Colwell & Lees 2000, Bakma et al. 2004). Geometric constraints are imposed on species geographic ranges by hard boundaries, i.e. boundaries acting as dispersal barriers (Colwell & Hurtt 1994). Basically, when species ranges are randomly placed in a domain, a peak of species richness emerges at the midpoint of this domain and the richness pattern is parabolic-shaped, without any direct effect from environmental, historical or biological factors (Colwell & Hurtt 1994, Colwell & Lees 2000). In this sense the MDE theory is very parsimonious and, as such, is considered a useful ‘null model’ for diversity patterns despite controversy about its main assumptions (Hawkins & Diniz Filho 2002, Colwell et al. 2004).

The MDE was originally described by Colwell & Hurtt (1994) within the context of latitudinal gradients in species diversity. A variety of null models have been proposed to test predictions of the MDE in 1 dimension. Indeed, MDE studies have focused on latitudinal (Willig & Lyons 1998, Koleff & Gaston 2001, Romdal et al. 2005), elevational (Sanchez Cordero 2001, Sanders 2002, McCain 2004) and bathymetric (Pineda & Caswell 1998, Kendall & Haedrich 2006) gradients (see Colwell et al. 2004 for a review). However, since species are distributed along longitudinal as well as latitudinal gradients, models were also developed for 2-dimensional species distributions (Bokma et al. 2001, Jetz & Rahbek 2001, Diniz Filho et al. 2002, Hawkins & Diniz Filho 2002, Rangel & Diniz Filho 2004). Overall, the evidence for MDE is contentious (see Colwell et al. 2004 for a review) with some studies claiming to find geometric constraints on species richness patterns (e.g. Willig & Lyons 1998, Jetz & Rahbek 2001) while others have failed to do so (e.g. Hawkins & Diniz Filho 2002, Kerr et al. 2006). Finally, Currie & Kerr (2008) proposed an analytic review of the MDE literature and concluded a general lack of support existed for the MDE theory.

Compared with patterns in other taxa, fish diversity patterns in terms of MDE have been poorly investigated (but see Bellwood et al. 2005, Kendall & Haedrich 2006), especially along longitudinal gradients (but see Hughes et al. 2002, Connolly et al. 2003). We observe that the Mediterranean Sea is ‘the archetypal situation’ where one would expect to observe a MDE similar to that described by Kerr et al. (2006) for Madagascar. Indeed, the Mediterranean Sea is a semi-enclosed area containing a large number of endemic fish species, and is bounded by 3 continents (Europe,
In this study we present an analysis of fish diversity patterns in the Mediterranean Sea and their potential determinants. Towards this objective we built the first comprehensive database on the distribution of 619 Mediterranean Sea fishes. We carried out 1-dimensional (latitudinal and longitudinal gradients) as well as 2-dimensional MDE analyses. In 2 dimensions, we found that by using (1) the MDE as a null model for fish richness distribution and (2) modelling techniques that take into account the spatial autocorrelation structure in the data, we obtained conclusions consistent with classical theories when investigating the potential influence of predators (environmental and historical factors such as primary production, temperature and distance from diversity sources) on Mediterranean fish diversity.

MATERIALS AND METHODS

Collected database and species pools. We created a database of the geographical distribution areas of all known fish species in the Mediterranean Sea (including endemic and exotic species) using GIS software (ArcView 3.3, ESRI). Data about exotic species were compiled by updating the list of the Mediterranean Science Commission (CIESM) Atlas (Golani et al. 2002) and the list of Quignard & Tomasini (2000). For all other species, data were compiled from the Fishes of the Northern Atlantic and Mediterranean (FNAM) atlas (Whitehead et al. 1986). This atlas is based on regional data sets and expert knowledge and still provides the best and the only available geographic ranges of all the fish species on a Mediterranean Sea scale. Recent field surveys (MEDITs project) were carried out only on the continental shelf and concerned only the northern Mediterranean Sea and a list of target groundfish species (Gaertner et al. 2007). We used the FNAM database to map the geographical distribution of fish richness by overlaying range maps of the 619 species.

According to Colwell & Hurtt (1994) and Jetz & Rahbek (2001) the mid-domain effect is more likely to emerge for endemic taxa, so we compared model predictions for 3 categories of species: the entire species pool, endemic fishes only and the entire pool except exotic fishes. It has also been argued that species with large ranges should be more influenced by geometric constraints than by species with smaller ranges (Colwell & Lees 2000, Jetz & Rahbek 2001, Dunn et al. 2007). Thus, we classified each species into one of 3 range classes as described by Hawkins & Diniz Filho (2002): (1) species with small ranges (<25% of the latitudinal extent of the Mediterranean Sea for the 1-dimensional longitudinal model and <25% of the longitudinal extent of the Mediterranean Sea for the 1-dimensional latitudinal model); (2) species with intermediate ranges (between 25 and 50% of the longitudinal extent of the Mediterranean Sea for the 1-dimensional longitudinal model and between 25 and 50% of the latitudinal extent of the Mediterranean Sea for the 1-dimensional latitudinal model); and (3) species with large ranges (>50% of the longitudinal extent of the Mediterranean Sea for the 1-dimensional longitudinal model and >50% of the latitudinal extent of the Mediterranean Sea for the 1-dimensional latitudinal model). We compared model predictions for each class separately as well as for all species considered together.

MDE models. One-dimensional MDE models: Empirical patterns of species richness in 1 dimension (longitudinal then latitudinal gradient) were compared with the predictions using 2 MDE models: the Colwell (2006) model and the Willig & Lyon (1998) model. First, we used Colwell's Range Model software v5 (Colwell 2006) as proposed by Dunn et al. (2006) to generate empirical species distributions in 1 dimension. Indeed, the new discrete model of Dunn et al. (2006) takes into account the gaps in the spatial distribution of species because ignoring these gaps would inflate species diversity estimates in the middle of the domain, which may amplify or create a mid-domain peak (Grytnes & Vetaas 2002). We called this MDE model CRMD after Colwell's (2006) Range Model as proposed by Dunn et al. (2006).

The Mediterranean Sea was divided into 423 longitudinal bins of 0.1° longitude and 156 latitudinal bins of 0.1° latitude. A species’ longitudinal (or latitudinal) range was measured as the number of bins (of 0.1° each) between the maximum longitude (or latitude) at which a species occurred and the minimum longitude (or latitude) at which a species occurred, including those extreme longitudes (or latitudes), regardless of how many occupied sites or unoccupied sites (gaps) occurred within the range. The ‘Fill’ parameter of the model was measured as the total number of bins in which a species occurred, including the extreme bins defining its range (Dunn et al. 2006).

The MDE model generated species richness patterns over the bins (the domain) that would be expected...
if observed species’ ranges were placed at random within the domain. Hence, the model maintains all observed species range sizes and occupancies while constraining ranges within domain limits (Dunn et al. 2006). ‘Range’ and ‘Fill’ parameters for each of the 619 species were imported into Colwell’s (2006) Range Model software and randomized 1000 times.

Secondly, we used the model of Willig & Lyons (1998). In the 1-dimensional (latitudinal) version of this null model, species richness at a point $P$ (a bin of 0.1° in our case) is a function of its distance to northern and southern endpoints of the domain, defined by the proportions $p$ and $q$. Species richness in $P$ is then given by $2pqS$, where $S$ is the total species richness in the pool. Using this model, that we called the WL model (Willig & Lyons 1998), we tested the relationship between empirical richness and $2pqS$. We also extended the latitudinal model to a longitudinal one where species richness at a point $P$ is a function of its distance to eastern and western endpoints of the domain, defined by the proportions $r$ and $t$. Then, predicted species richness in $P$ is given by $2rtS$.

**Two-dimensional MDE model:** Given that the distribution of species on a domain occurs in 2 dimensions we also explored fish richness patterns using the null model of Willig & Lyons (1998) extended by Bokma et al. (2001). Species richness at point $P$ (a cell of 0.1° by 0.1° in our case) becomes a function of its distance to northern, southern, eastern and western endpoints of the domain, i.e. $4pqrtS$. For this, we used a regular grid with a total of 27 019 cells over the Mediterranean Sea (0.1° latitude × 0.1° longitude). For each cell, the values of $p$, $q$, $r$ and $t$ were calculated by considering their positions relative to the northern, southern, eastern and western endpoints of the whole Mediterranean Sea, respectively, and, thus, to the absolute continental boundary. Here, we did not consider the positions relative to the immediate continental boundary in order to reduce the diversity peaks in bays, gulfs and inlets. Indeed, as mentioned by Hawkins & Diniz Filho (2002), if we consider the immediate boundary, semi-enclosed areas such as gulfs would have a high predicted richness because $p$, $q$, $r$ and $t$ proportions will be set to their borders and not to the entire domain.

More than 75% of Mediterranean fish species are strictly coastal, i.e. do not live beyond the limits of the continental shelf. We, thus, extracted the set of 8005 cells corresponding to the limit of 200 m water depth from the 27 019 cells constituting the whole Mediterranean Sea dataset. We carried out 2-dimensional MDE analyses on this subset as well.

**MDE predictive power assessment:** The predictive power of the MDE can be assessed using the coefficient of regression $R^2$ between observed and simulated species richness values (e.g. Diniz Filho et al. 2002, Rangel & Diniz Filho 2004). However, the adequate test of these regressions needs to take into account the spatial dependence among data (bins or cells). In the majority of macroecological studies dealing with species richness patterns, observations from nearby locations tend to be more similar than observations randomly chosen: observations are spatially autocorrelated sensu Legendre (1993). As a consequence, using the number of bins or cells to determine degrees of freedom for inference is not valid and may produce spurious rejections of the null hypothesis. As an alternative, we adjusted the number of degrees of freedom following Dutilleul (1993): basically the correction takes into account the observed pattern in spatial autocorrelation to calculate the geographically effective number of degrees of freedom. This method was implemented using SAM 3.0 software (Spatial Analysis in Macroecology, Rangel et al. 2006).

**Regression modelling. Predictor variables:** Sea surface temperature (SST) is known to be highly heterogeneous in the Mediterranean Sea with the eastern basin being warmer than the western one and these differences may limit fish dispersion (e.g. Ben Rais Lasram et al. 2008). More generally, climate constrains species richness across taxa at large scales (Currie et al. 2004). Hence, we included annual mean, maximum and minimum SST as predictor variables in our models. First, we collected a grid (1°) of weekly SST values from the National Climatic Data Center (NCDC) National Operational Model Archive and Distribution System Meteorological Data Server (NOMADS) of the US National Oceanic and Atmospheric Administration (NOAA) Satellite and Information Service (www.osdps.noaa.gov/PSB/EPS/SST/al_climo_mon.html). Then, we interpolated maps at the 0.1° resolution needed for our analyses via ordinary kriging (Diggle & Ribeiro 2007). Weekly SST values were averaged over the 1998 to 2002 period to provide 14 SST variables for each 0.1° grid cell. Twelve monthly means were used to estimate the mean annual SST while an absolute minimum and an absolute maximum were extracted for that period.

The Mediterranean Sea basin is also characterised by a strong gradient of primary production decreasing from the richer northwestern Mediterranean Sea (350 to 450 mgC m$^{-2}$ d$^{-1}$, Moutin & Raimbault 2002) to the more oligotrophic eastern basin (150 mgC m$^{-2}$ d$^{-1}$, Turley et al. 2000). Since energy supply towards upper trophic levels is determined by primary production, the energy-richness hypothesis posits that productive energy is one of the most important factors driving species richness across taxonomic groups (Mittelbach et al. 2001). In the Mediterranean Sea we considered the surface chlorophyll concentration as a measure of primary production. Production data were provided by Bosc et al. (2004) who used the ocean color sensor Seaviewing
Wide Field-of-view Sensor (SeaWiFS) to estimate surface chlorophyll concentration in the sea’s upper layer. The computation was based on predictions from the model of Antoine & Morel (1996) averaged over the 1998 to 2002 time period. As we did for SST, primary production data were interpolated via ordinary kriging (Diggle & Ribeiro 2007) to generate a primary production map corresponding to our 0.1° grid.

Similarly, as with most macroecological studies, historical contingencies are likely to influence our fish diversity patterns. Indeed, the re-opening of the Strait of Gibraltar at the dawn of the Pliocene Epoch (5 million years ago) as well as the completion of the Suez Canal in 1869, have been suspected to play a major role in the diversity distribution of the Mediterranean Sea (e.g. Gaertner et al. 2007, Danovaro et al. 2008). Consequently, we introduced the distance to these 2 diversity sources as explanatory variables. Those distances were calculated using the GIS software ArcView 3.3. For each cell, the distance to the immediate continental boundary was measured, and then added to the distance separating this continental boundary from the Strait of Gibraltar and the Suez Canal along the coastline.

**Non-spatial modelling:** Given the current debate about their general relevance (e.g. Colwell et al. 2004), MDE predictions may or may not be included in the construction of multi-predictor models for diversity patterns. If the MDE is not considered, the usual ordinary least squares (OLS) multi-predictor regression model is:

\[ Y_{\text{obs}} = X\beta + \varepsilon \] (1)

where \( Y_{\text{obs}} \) is the observed species richness, \( \beta \) is the vector of coefficients associated to the predictors \( X \) and \( \varepsilon \) represents the (spatially) independent and identically distributed residuals.

When the MDE is considered, its predictions can be incorporated in one of the following ways (Currie & Kerr 2008):

1. The MDE is considered as a null model and its predictions can be removed from observed patterns before testing for environmental effects (e.g. Connolly et al. 2003). This is achieved through the modelling of a new ‘deviation from MDE’ variable \( Y_{D} \):

\[ Y_{D} = Y_{\text{obs}} - Y_{\text{MDE}} \] (2)

where \( Y_{\text{MDE}} \) is the predicted richness according to the mid-domain hypothesis. The multi-predictor regression model is thus:

\[ Y_{D} = X\beta + \varepsilon \] (3)

2. The MDE is used as a predictor like an environmental or historical variable. The relative importance of each of these predictors is examined in the same model (e.g. Bellwood et al. 2005):

\[ Y_{\text{obs}} = X\beta + \delta Y_{\text{MDE}} + \varepsilon \] (4)

where \( \delta \) is the coefficient associated with MDE predictions (e.g. Bellwood et al. 2005).

Since the Mediterranean Sea represents the archetypal situation where one would expect to observe the MDE and to avoid problematic interpretation of model coefficients linked to the potential collinearity of MDE with other predictors (Currie & Kerr 2008), we chose the model described in Eq. (3). We considered the MDE as a null model and built a multi-predictor model to explain the deviations from MDE predictions using environmental and historical factors.

**Spatial modelling:** Given that the presence of spatial autocorrelation in the data is problematic for statistical analyses such as regressions, which assume independently distributed errors, we investigated whether the use of spatial regression models should improve the analysis. More precisely, we used simultaneous autoregressive (SAR) models that incorporate the spatial autocorrelation structure among observations. Following the comparative analysis performed by Kissling & Carl (2008), we implemented the spatial error model (SARerr model hereafter). The SARerr model complements the OLS model with a term \( \Lambda W\omega \) designed to capture both ‘inherent’ and ‘induced’ spatial autocorrelation (Kissling & Carl 2008) for species richness or alternatively for MDE deviations:

\[ Y_{\text{obs}} = X\beta + \Lambda W\omega + \varepsilon \] (5)

\[ Y_{D} = X\beta + \Lambda W\omega + \varepsilon \] (6)

where \( \Lambda \) is the spatial autoregression coefficient, \( W \) is the spatial weights matrix indicating the spatial relationships among observations (expressed as species richness or deviations from the MDE predictions for each cell) and \( \omega \) is the spatially dependent error term.

We calculated the spatial weights matrix based on a neighbourhood graph.

Spatial modelling was carried out using the library ‘spdep’ for the R statistical programming environment (R Development Core Team 2007).

**Predictor selection and model evaluation:** We implemented non-spatial OLS and SARerr models for both observed richness (\( Y_{\text{obs}}, \) Eqs. 1 & 5) and MDE deviations (\( Y_{D}, \) Eqs. 3 & 6) using 3 different species pools, resulting in 12 regression models.

For non-spatial regression models, the predictor selection was based on a backward selection procedure starting with a model including all predictors. The removal of predictor variables was based on chi-squared deletion tests (Crawley 2007). The minimal adequate model was selected as the one containing nothing but significant predictors. To assess the robustness of our results to the model selection procedure, we complemented the chi-squared procedure with a backward,
stepwise, Akaike’s Information Criterion (AIC) model selection algorithm. Since both procedures resulted in the selection of the complete model in all cases, we do not comment on the comparison again.

Models were evaluated based on the percentage of explanation ($R^2$) provided by predictors on fish richness (or deviation from MDE predictions). We compared the goodness of fit of different models in relation to their complexity using the AIC, which incorporates fit quality while penalizing for complexity (Burnham & Anderson 2002). To investigate the spatial dependence of model residuals, we tested whether the observed Moran’s index ($I$) was statistically different from those expected by chance using a permutation test (‘spdep’ package, R software, R Development Core Team 2007).

**RESULTS**

**Fish diversity patterns in the Mediterranean Sea**

The first maps representing fish species richness as well as endemic fish richness at the Mediterranean Sea scale are provided in Fig. 1. Species richness exhibits a decreasing gradient from west to east as claimed by previous studies (e.g. Quignard & Tomasini 2000). Sicily is the hotspot of highest richness with 375 species per cell (Fig. 1a). The endemic richness gradient is more pronounced latitudinally, i.e. from north to south: the northern side exhibits a greater species richness and the Adriatic Sea appears as a hotspot of endemism with 45 species per cell (Fig. 1b).

**Geographic range size distributions**

Longitudinally, the entire pool of fish species contained 153 small, 68 intermediate and 398 large range species. The 79 endemic species were classified into 27 small, 12 intermediate and 40 large range species. The ‘without exotics’ pool comprised 68 small, 45 intermediate and 379 large range species. Latitudinally, the 619 species of the entire pool were classified into 112 small, 95 intermediate and 412 large range species. The endemic pool contained 20 small, 19 intermediate and 40 large range species. The ‘without exotics’ pool consisted of 49 small, 54 intermediate and 389 large range species.

![Fig. 1. Observed species richness gradient for the (a) entire and (b) endemic species pools. Spa. = Spain, Fra. = France, Ita. = Italy, Gre. = Greece, Tur. = Turkey, Isr. = Israel, Egy. = Egypt, Lib. = Libya, Tun. = Tunisia, Alg. = Algeria, Mor. = Morocco](image-url)
One-dimensional analyses

Observations and MDE predictions (CRMD and WL models) of fish species richness for the 3 species pools are presented for longitudinal and latitudinal gradients in Fig. 2 (see Figs. A1 to A12 in Appendix 1 for observations and predictions corresponding to different range categories; available as supplementary material at http://www.int-res.com/articles/suppl/m376p253_app.pdf). Results of correlation analysis between observed and predicted richness patterns are presented in Tables 1 & 2 for longitudinal and latitudinal gradients, respectively. Degrees of freedom (df) and p-values took into account the correction for spatial autocorrelation.

Longitudinally, fish richness patterns were highly asymmetric for the entire species pool and for the ‘without exotics’ pool, and, thus, differed markedly from MDE predictions ($R^2 < 0.45$ and $p > 0.05$ in Table 1) (Fig. 2a). However, the longitudinal gradient of endemic fish richness appeared much closer to MDE predictions. Indeed, the strongest and significant coefficient of determination between simulated and observed fish richness was obtained for endemic species ($R^2 = 0.68$ for the CRMD model and $R^2 = 0.71$ for the WL model with $p < 0.05$, Table 1). Endemic species richness exhibited a pronounced mid-longitude peak consistent with predictions of a simple 1-dimensional MDE model between 13.45° and 15.55° E longitude, which corresponds to that in the middle of the Mediter-

![Fig. 2. Observed and predicted species richness curves for the entire species pool, the endemic species pool and the 'without exotic' species pool along (a) longitudinal and (b) latitudinal gradients according to Colwell's (2006) model as proposed by Dunn et al. (CRMD, 2006) and to Willig & Lyon’s (1998) model (WL). Circles indicate the observed richness, solid and dashed lines indicate the predicted richness according to CRMD and to WL models, respectively](image)
When endemic species were split among the 3 range categories, a strong MDE was observed for large range species ($R^2 = 0.79$ for both the CRMD and WL models with $p < 0.05$, Table 1). The WL model also revealed a significant MDE for endemic species with intermediate range ($R^2 = 0.55$, $p < 0.05$). However, for the small-ranged species, the MDE did not significantly explain endemic richness patterns ($R^2 = 0.23$ for both models, Table 1).

### Two-dimensional analyses

The model of Willig & Lyons (1998) extended by Bokma et al. (2001) produced a peak of fish richness in the middle of the Mediterranean Sea where the value of $4\sqrt{pqrstS}$ is the highest, while a diversity of zero was predicted at the extremities of the domain where 1 of the 4 parameters ($p, q, r$ or $s$) is 0 (Fig. 3). The peak of predicted fish richness was produced along the Greek, south Italian, north Tunisian and Sardinian coasts, as well as in the middle of the water body separating these coasts (Fig. 3).

On the continental shelf, the relationship between observed and predicted (2-dimensional MDE) fish richness values was only significant for the pool ‘without exotics’ (Table 2). When the entire pool was split into 3 range categories, a strong MDE was observed for large range species ($R^2 = 0.68$ for both the CRMD model and $R^2 = 0.69$ for the WL model with $p < 0.05$, Table 2). It was also the case for the intermediate range species of the ‘without exotics’ pool where $74\%$ (according to CRMD model with $p < 0.05$) and $72\%$ (according to WL model with $p < 0.05$) of the variation in species richness was explained by MDE predictions (Table 2).

### Table 1. Results of the regressions between observed and predicted fish species richness along the longitudinal gradient using the mid-domain effect (MDE) null models of Colwell (2006) (CRMD) and Willig & Lyons (1998) (WL). $n =$ number of species, $df =$ degrees of freedom, $R^2 =$ coefficient of determination. *$p < 0.05$, other values are not significant

<table>
<thead>
<tr>
<th>Species pool</th>
<th>Range class</th>
<th>n</th>
<th>CRMD df</th>
<th>CRMD $R^2$</th>
<th>WL df</th>
<th>WL $R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Entire</td>
<td>All</td>
<td>619</td>
<td>2.2</td>
<td>0.24</td>
<td>2.1</td>
<td>0.24</td>
</tr>
<tr>
<td></td>
<td>Small</td>
<td>68</td>
<td>2.9</td>
<td>0.07</td>
<td>6.9</td>
<td>0.08</td>
</tr>
<tr>
<td></td>
<td>Intermediate</td>
<td>153</td>
<td>1.7</td>
<td>0.15</td>
<td>2</td>
<td>0.15</td>
</tr>
<tr>
<td></td>
<td>Large</td>
<td>398</td>
<td>2.3</td>
<td>0.45</td>
<td>2.3</td>
<td>0.36</td>
</tr>
<tr>
<td>Endemic</td>
<td>All</td>
<td>79</td>
<td>1.7</td>
<td>0.68*</td>
<td>1.6</td>
<td>0.71*</td>
</tr>
<tr>
<td></td>
<td>Small</td>
<td>27</td>
<td>12</td>
<td>0.23</td>
<td>4.7</td>
<td>0.23</td>
</tr>
<tr>
<td></td>
<td>Intermediate</td>
<td>12</td>
<td>1.6</td>
<td>0.60</td>
<td>36.3</td>
<td>0.55*</td>
</tr>
<tr>
<td></td>
<td>Large</td>
<td>40</td>
<td>1.7</td>
<td>0.79*</td>
<td>1.5</td>
<td>0.79*</td>
</tr>
<tr>
<td>Without exotics</td>
<td>All</td>
<td>492</td>
<td>2.3</td>
<td>0.27</td>
<td>2.2</td>
<td>0.27</td>
</tr>
<tr>
<td></td>
<td>Small</td>
<td>68</td>
<td>7.1</td>
<td>0.05</td>
<td>5.8</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>Intermediate</td>
<td>45</td>
<td>1.7</td>
<td>0.15</td>
<td>1.9</td>
<td>0.14</td>
</tr>
<tr>
<td></td>
<td>Large</td>
<td>379</td>
<td>2.6</td>
<td>0.33</td>
<td>2.4</td>
<td>0.34</td>
</tr>
</tbody>
</table>

### Table 2. Results of the regressions between observed and predicted fish species richness along the latitudinal gradient using the mid-domain effect (MDE) null models of Colwell (2006) (CRMD) and Willig & Lyons (1998) (WL). $n =$ number of species, $df =$ degrees of freedom, $R^2 =$ coefficient of determination. *$p < 0.05$, other values are not significant

<table>
<thead>
<tr>
<th>Species pool</th>
<th>Range class</th>
<th>n</th>
<th>CRMD df</th>
<th>CRMD $R^2$</th>
<th>WL df</th>
<th>WL $R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Entire</td>
<td>All</td>
<td>619</td>
<td>6.1</td>
<td>0.80*</td>
<td>1.9</td>
<td>0.80*</td>
</tr>
<tr>
<td></td>
<td>Small</td>
<td>112</td>
<td>9.3</td>
<td>0.14</td>
<td>5.7</td>
<td>0.12</td>
</tr>
<tr>
<td></td>
<td>Intermediate</td>
<td>95</td>
<td>1.7</td>
<td>0.54</td>
<td>1.9</td>
<td>0.54</td>
</tr>
<tr>
<td></td>
<td>Large</td>
<td>412</td>
<td>2.2</td>
<td>0.68*</td>
<td>1.8</td>
<td>0.69*</td>
</tr>
<tr>
<td>Endemic</td>
<td>All</td>
<td>79</td>
<td>6.9</td>
<td>0.13</td>
<td>2.8</td>
<td>0.15</td>
</tr>
<tr>
<td></td>
<td>Small</td>
<td>20</td>
<td>18.6</td>
<td>0.06</td>
<td>59.6</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>Intermediate</td>
<td>19</td>
<td>2</td>
<td>0.11</td>
<td>2.2</td>
<td>0.12</td>
</tr>
<tr>
<td></td>
<td>Large</td>
<td>40</td>
<td>3.2</td>
<td>0.16</td>
<td>3</td>
<td>0.18</td>
</tr>
<tr>
<td>Without exotics</td>
<td>All</td>
<td>492</td>
<td>6</td>
<td>0.63*</td>
<td>1.8</td>
<td>0.65*</td>
</tr>
<tr>
<td></td>
<td>Small</td>
<td>49</td>
<td>9.6</td>
<td>0.15</td>
<td>7.9</td>
<td>0.21</td>
</tr>
<tr>
<td></td>
<td>Intermediate</td>
<td>54</td>
<td>3.3</td>
<td>0.74*</td>
<td>1.5</td>
<td>0.72*</td>
</tr>
<tr>
<td></td>
<td>Large</td>
<td>389</td>
<td>9.6</td>
<td>0.15</td>
<td>3.3</td>
<td>0.14</td>
</tr>
</tbody>
</table>
Regression analyses

The results obtained for all regression models where the predicted variable was either fish species richness or fish richness deviations from MDE predictions are summarized in Table 3. All the non-spatial OLS models showed a strong spatial autocorrelation pattern in the residuals (significant Moran’s $I$) for both fish richness and MDE deviations as explained variables. The SAR$_{err}$ models successfully removed...
The Mediterranean Sea has been intensively studied since the 15th century (Bianchi & Morri 2000). More recently, various investigations have tried to inventory macroorganisms, especially fishes, in the Mediterranean Sea. The most famous ones are the Medifaune database (Fredj & Maurin 1987), the Species Identification Guide for Fishery Purposes of the UN Food and Agriculture Organization (FAO) (Fischer et al. 1987) and the FNAM atlas (Whitehead et al. 1986) that we used in this study. This atlas still provides the only available geographic ranges for all Mediterranean fish species and is based on regional data and expert knowledge. However, some areas of the southern Mediterranean Sea suffer from a lack of scientific surveys, especially the Libyan coast, making the data somewhat heterogeneous. The non-detectability is problematic because the patterns can be biased, but this problem is inevitable and common to all studies carried out at the macro-scale. Even if we have to be cautious with the results we believe that fish diversity gradients are so marked that discovering occurrences of new species in some parts of the Mediterranean Sea would probably not change our conclusions.

The Mediterranean Sea can be defined as an archetypal domain where we expect a MDE because this biome is (1) a big domain, (2) an enclosed area with well-defined boundaries, (3) has a high rate of endemic species (8.8%) and (4) has a significant proportion of large range species. While a MDE was displayed in 1 dimension (in longitude for the endemic species pool and in latitude for the entire and the ‘without exotics’ pools), our data on fish richness were less consistent with the 2-dimensional patterns predicted by the null MDE model. Overall, our analyses supported an evident expectation of MDE in 1 dimension, namely that richness of endemic large-ranged species should be more consistent with MDE predictions than that of small-ranged species (e.g. Lees et al. 1999).

Secondly, our results indicated that diversity patterns of fish species in the Mediterranean Sea were also driven by historical or environmental factors. Indeed our models (Table 3) showed that environmental variables and historical factors explained a large proportion of species richness.

Table 3. Standardized coefficients of 6 explanatory variables used to model patterns of species richness of fishes with 2 modelling techniques: non-spatial ordinary least squares (OLS) and spatial autoregressive model (SAR). TMOY: mean sea surface temperature (SST), TMAX: maximum SST, TMIN: minimum SST, PROD: primary production, DGI: distance to the Strait of Gibraltar, DSU: distance to the Suez Canal, Pred: fish richness (FR) and fish richness deviation from MDE predictions (D), I: Moran’s index of residual autocorrelation, R² = coefficient of determination. *p < 0.05, other values are not significant. AIC values obtained by including the significant parameters only

<table>
<thead>
<tr>
<th>Model</th>
<th>Pred</th>
<th>Predictor variables</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>TMOY</td>
</tr>
<tr>
<td>Entire species pool</td>
<td>OLS</td>
<td>FR</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>-45.4*</td>
</tr>
<tr>
<td></td>
<td>SAR</td>
<td>FR</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>-37.9*</td>
</tr>
<tr>
<td>Endemic species pool</td>
<td>OLS</td>
<td>FR</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>-17*</td>
</tr>
<tr>
<td></td>
<td>SAR</td>
<td>FR</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>-12.8*</td>
</tr>
<tr>
<td>'Without exotics' species pool</td>
<td>OLS</td>
<td>FR</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>-55.3*</td>
</tr>
<tr>
<td></td>
<td>SAR</td>
<td>FR</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>-49.3*</td>
</tr>
</tbody>
</table>
Although absolute deviations were more important for the WL model, 1-dimensional null models provided approximately the same coefficients of determination between observed and predicted fish richness patterns (Tables 1 & 2). The average difference between the coefficients of determination of the 2 models was only 1% longitudinally and 0.5% latitudinally.

One-dimensional patterns and historical contingencies

Longitudinally, the MDE had a high explanatory power on fish diversity patterns for the pool of endemic species with large ranges (79% for both CRMD and WL models, p < 0.05). Indeed, geometric constraints forced large-ranged endemic species to occupy the middle of the Mediterranean Sea where they increased fish species richness. This is a true MDE. The MDE was still significant for the entire endemic pool. Indeed, most of the species with small ranges were constrained to the Adriatic Sea situated in the middle of the Mediterranean Sea (between 12.25° and 19.50° E). Consequently, in addition to the peak produced by large-ranged species, a peak produced by small-ranged species situated in the bins including the Adriatic Sea produced a pattern consistent with MDE predictions.

When we examined the latitudinal pattern, the MDE was not significant for the endemic species pool (Table 2). This is due to the latitudinal position of the Adriatic Sea, which concentrates many endemic species and produces a strong deviation from latitudinal MDE predictions. Conversely, when we considered the entire and the ‘without exotics’ pools we observed a strong latitudinal MDE. This pattern may be explained by the high species richness observed along the Sicilian coasts that are included in the latitudinal central bins of the Mediterranean Sea, i.e. between 37.85° and 38.15° N. In this bin interval, deviations from the MDE model were close to zero (Fig. 5). This pattern may be explained by the position of Sicily, which acts as a barrier (physical and thermal) where fishes from the 2 basins meet and produce species rich assemblages.

Many studies have claimed the existence of a clear gradient from east to west: species richness decreases from the western Mediterranean Sea basin to the eastern one (Quignard & Tomasini 2000, Taviani 2002, Emig & Geistdoerfer 2004). The most classical explanatory factors are the threshold materialised by the Siculo-Tunisian Strait that divides the Mediterranean Sea into 2 basins and the paleo-biogeographical history of the Mediterranean Sea. Indeed, during the Cretaceous Period, the Mediterranean Sea (called Thetys Sea at that stage) was connected to the Atlantic and the Indo-Pacific oceans (Bianchi & Morri 2000). Both oceans filled the Thetys Sea with very different faunas. During the Miocene Epoch, the Thetys Sea was isolated from the Indo-Pacific Ocean and at the Messinian stage, the connection with the Atlantic Ocean closed. During this Messinian salinity crisis, the Mediterranean underwent a severe desiccation that drove...
many species to extinction. However, some shallow areas remained on the 2 sides of the Siculo-Tunisian Strait, which allowed many allopatric speciations to occur (Bianchi & Morri 2000, Heads 2005, Rouchy & Caruso 2006). At the reopening of the Strait of Gibraltar 5 million years ago, the Atlantic Ocean again provided the Mediterranean Sea with fauna and flora. Consequently, the western basin, which is cooler and richer, shows more biological similarity with the Atlantic Ocean while the eastern basin shows more biological similarity with the Indo-Pacific Ocean especially after the Suez Canal opened in 1869 by the architect Ferdinand de Lesseps (Bianchi & Morri 2000), thereby allowing colonization by Lessepsian species to occur (e.g. Ben Rais Lasram et al. 2008). The Siculo-Tunisian Strait still acts as a barrier to the dispersal of many species between the 2 basins and at the same time constitutes their meeting point. Hence, since this barrier is likely to produce a fish richness peak in a central point of the Mediterranean Sea, we obtained a MDE signal. However, this barrier may have been established closer to Mediterranean boundaries and may have produced 2 basins of different sizes. In this case the MDE signal would have been absent while the colonization process would have been the same suggesting that historical contingencies may or may not produce species richness peaks according to MDE predictions.

**Particularities when considering exotic species**

When we test the MDE, we may exclude exotic species from the dataset because their geographic ranges are less likely to be drawn randomly (Colwell et al. 2004). This was particularly false for the Mediterranean Sea. Our entire pool contained 127 exotic species (65 from the Suez Canal and 62 from the Strait of Gibraltar). Because exotic species were introduced on opposite sides of the Mediterranean Sea, the MDE may be expected to be distorted because these introductions would increase species richness at the 2 extremities of the domain. This assumption was true longitudinally, but the change observed in $R^2$ was low when we removed exotic species (from $R^2 = 0.24$ to $R^2 = 0.27$). However, the contribution of the MDE to the latitudinal richness gradient decreased when we removed exotic species (Table 2). Indeed, exotic species spread from the outflows of invasion sources (Suez Canal and Strait of Gibraltar) and spread preferentially along the Turkish and Greek coasts in the eastern basin and along the Spanish coasts in the western basin (Ben Rais Lasram et al. 2008). Since these areas are situated within the central latitudinal bins in the Mediterranean Sea, removing them would decrease species richness in the middle of the domain and the observed richness would depart from expectation under MDE. Moreover, introduced species from the Atlantic Ocean and from the Red Sea, which spread in the Mediterranean Sea, are likely to overlap at the intersection of the 2 basins, consequently increasing the MDE.

This particularly induced the opposite pattern in 2 dimensions. Indeed, when we deal with cells and not with bins, exotic species distort the MDE. Removing them makes species richness values closer to MDE predictions: variations in species richness explained by MDE predictions in 2 dimensions were low for the entire pool, but not for the ‘without exotics’ pool where MDE explained 21% of variations in species richness ($p < 0.05$). Contribution of the MDE was, therefore, not negligible for that pool: the highest richness values were found around Sicily, northern Tunisia and Sardinia, which was consistent with predictions of the 2-dimensional MDE model (Fig. 4).

**Two-dimensional patterns and deviations from a mid-domain model**

Even if the MDE accounted for a small part of variation in fish richness patterns, our results, using regression modelling (Table 3), pointed out that sequentially taking into account the MDE and the spatial autocorrelation in assessing the effect of environmental and historical factors had the potential to dramatically modify the conclusions that can be drawn. Accordingly, as with most macroecological studies (Currie et al. 2004), we showed that climate–richness relationships had the strongest effect on Mediterranean fish diversity. However, these conclusions cannot be drawn from a direct analysis of observed patterns with the use of simple tools such as non-spatial OLS regressions. Indeed, by analysing the raw data directly, we saw that minimum temperature and primary production negatively influenced fish richness, which is counterintuitive and in contrast to the species–energy hypothesis. Conversely and intuitively, performing regressions on deviations from the MDE predictions, these 2 variables positively affected the Mediterranean fish diversity. Furthermore, when accounting for the spatial autocorrelation in the data, the significant influence of the distances to diversity sources on diversity patterns were challenged. The reason for such discrepancies in the results of statistical analyses certainly lies in the fact that, even under any environmental effect, the expected fish richness pattern in the Mediterranean Sea is not homogeneous (the same fish richness in each cell), but instead would follow a MDE pattern subjected to spatial autocorrelation. When we took into account the MDE and the spatially autocorrelated nature of observations (SAR$_{err}$ models for deviation from MDE predic-
tions) we found that fish species richness was positively influenced by primary production and minimum SST, while maximum and mean SST had negative influences on fish richness. This pattern was mainly driven by the species rich assemblages of the northwestern Mediterranean Sea (much richer than predicted by the MDE) associated with cooler mean SST and high primary production. This result also suggested that water temperature is a strong limiting factor for many species with metabolism limited by low temperature (Ibarz et al. 2007), while high temperature may lead to size limitation (Sonin et al. 2007).

**CONCLUSION**

Overall, the results of our analyses in the Mediterranean Sea suggested: (1) in 1 dimension the only and true MDE is displayed longitudinally for the large-ranged endemic species while the ‘apparent’ MDE for other species pools is related to historical and paleo-geographical contingencies; and (2) in 2 dimensions, analyzing the deviations from MDE predictions using autoregressive models allowed a more rigorous assessment of the factors influencing fish diversity patterns. Even if the MDE accounted for a non-negligible proportion of species richness variation, it was not sufficient to explain the fish diversity pattern in the Mediterranean Sea, which corroborates other bi-dimensional studies testing the MDE (Bokma et al. 2001, Jetz & Rahbek 2001, Diniz Filho et al. 2002, Hawkins & Diniz Filho 2002, Rangel & Diniz Filho 2004). Other processes such as historical and environmental factors should be considered. Using regression models that take into account spatial autocorrelation, we found the most common predictions of macroecology; that is, there is a positive influence of primary production and minimum temperature on fish species richness.

Since biodiversity assessment is moving towards a multifaceted framework including phylogenetic and functional differences among species (Floeter et al. 2008, Mouillot et al. 2008, Villeger et al. 2008), a new challenge for marine ecology is to investigate the concordance between environmental and historical determinants of functional, phylogenetic and taxonomic diversities.

**Acknowledgements.** The authors express their gratitude to the Cooperation and Cultural Action Services of the French Embassy in Tunisia, which funded this research with a PhD grant. They also thank E. Bosc (Marine Environment Laboratory at the International Atomic Energy Agency, Monaco) who kindly provided the primary production data. This project was also supported by the Total Foundation. Four anonymous reviewers provided constructive comments on a first draft of the manuscript.

**LITERATURE CITED**


Editorial responsibility: Lisandro Benedetti-Cecchi, Pisa, Italy

Submitted: February 6, 2008; Accepted: October 21, 2008
Proofs received from author(s): January 30, 2009