ABSTRACT

Aim The study of areas of sympatry of species with predominantly parapatric distributions can provide valuable insights into their evolutionary history and the factors shaping patterns of species co-occurrence. This information is key in biogeography, evolutionary biology and conservation planning. In this study we analyse the distributions of two pairs of partially co-occurring congeneric amphibian species: tree frogs (Hyla molleri and H. meridionalis) and midwife toads (Alytes obstetricans and A. cisternasii).

Location Iberian Peninsula (SW Europe).

Methods We obtained distribution data from the herpetological atlases of Portugal and Spain, consisting of presences and absences on UTM 10 × 10 km grid cells. We built an environmental favourability model for each species, using 24 potential predictor variables representative of physiography, climate and human activity. Variables were selected for each model using both information and significance criteria. Models were evaluated using both calibration and discrimination measures. Models were then combined using fuzzy intersection, and compared using correlation analysis (accounting for spatial autocorrelation), niche comparison metrics and fuzzy similarity indices.

Results We found significant dissimilarity between the favourability patterns for A. obstetricans and A. cisternasii, indicating environmental segregation of these two midwife toad species. In tree frogs, we found significant similarity between favourability for H. meridionalis and for its co-occurrence with H. molleri – that is, sympatry occurs mainly in areas that are favourable for H. meridionalis.

Main conclusions These results provide clues to understand the evolutionary history of these four species, including the evolution of reproductive isolation, and suggest that conservation efforts for tree frogs may be focused on the areas that are favourable for both species, whereas midwife toads will require specific measures tailored for each species.

Keywords amphibians, distributional overlap, environmental favourability, evolutionary biogeography, fuzzy logic, niche evolution, parapatry, species distribution modelling, sympatry

INTRODUCTION

Many species share common distribution ranges, presenting considerable geographical areas of co-occurrence. Studying spatial and environmental patterns of overlap can provide valuable insights into the factors that determine the co-occurrence of species and shed light on aspects of their evolutionary history (Coyne & Orr, 2004). For instance, the characterization of spatial patterns of environmental favourability in species with partially overlapping ranges can provide
clues about the geographical and ecological context of the speciation process (Fisher-Reid et al., 2013; Wooten et al., 2013). While ecological niche differentiation is not necessarily a pre-requisite for speciation, niche divergence can act as a pre-zygotic isolating mechanism in spatially continuous populations, reinforcing or consolidating reproductive isolation in nascent species (Nosil et al., 2009). On the other hand, a lack of differentiation may result from phylogenetic niche conservatism (i.e. a tendency of species to retain their ancestral traits), which has been shown to be an important feature in allopatric speciation across a variety of taxa (Wiens et al., 2010).

Based on this reasoning, we hypothesize that comparative analyses of the environmental favourability patterns of parapatric species can help to distinguish the dominant ecological/evolutionary forces shaping their distributions, in terms of independence, conservatism or divergence of niche (Fig. 1). We thus model and compare the distributions of two pairs of parapatric amphibian species in the Iberian Peninsula: tree frogs (Hyla molleri and H. meridionalis), with 11% of the occurrence area (based on UTM 10 × 10 km cells) shared by both species; and midwife toads (Alytes obstetricans and A. cisternasi), with an overlap of 4% in their occurrence areas in this region (Loureiro et al., 2010; MAGRAMA, 2015).

Species in each of these pairs differ in aspects of their natural history and ecological preferences at the local scale. For instance, in Alytes there seems to be little overlap in the breeding seasons of the two species (Malkmus, 2004), which limits the chances for hybridization. Also, A. cisternasi usually breeds in more temporary habitats, compared to the generally permanent water bodies selected by A. obstetricans. With respect to the terrestrial habitat, A. cisternasi usually occurs in open areas and avoids limestone or littoral dunes, whereas A. obstetricans can be found in a wider range of habitats, including dense forests, and in many types of geological substrates, including limestone or gypsum terrains (Garcia-París et al., 2004; Malkmus, 2004). There are also differences in their preferred elevation ranges (A. cisternasi populations are mostly found below 500 m, whereas A. obstetricans prefers places between 400 and 1000 m and can reach much higher, up to > 2000 m). Finally, A. cisternasi is present in meso- or thermomediterranean areas, with annual precipitation < 900 mm, whereas A. obstetricans is present in a wider variety of bioclimatic regions, but generally where annual precipitation is > 900 mm (Malkmus, 2004).

Conversely, the two Hyla species show fewer ecological differences. For instance, their terrestrial habitats are very similar (although H. meridionalis has a broader range of habitat types), and they seem to have synchronized their breeding seasons, thus probably increasing the chances for hybridization. Ecological differences among Hyla include their preferred elevation ranges (wider in H. molleri, but rarely > 500 m in H. meridionalis) and precipitation levels (high annual precipitation in H. molleri, but < 1000 mm in H. meridionalis) (Malkmus, 2004).

These local-scale ecological differences among Hyla and Alytes species were also identified in previous attempts to model their distributions (Arntzen, 2006) and suggest the existence of niche segregation. The two species pairs have similarly old divergence times (Miocene; see Smith et al., 2003; Maia-Carvalho et al., 2014), but the temporal scale of their overlap is markedly different: H. meridionalis is inferred to have entered the Iberian Peninsula in recent times, perhaps even in historical times, via human-mediated introduction (Recuero et al., 2007), whereas A. obstetricans and A. cisternasi have both evolved in situ and coexisted for millions of years (Martinez-Solano et al., 2004; Maia-Carvalho et al., 2014).

In amphibians, vicariance is the most common mode of speciation (Vences & Wake, 2007), and therefore the study of areas of secondary contact is key to assess patterns of reproductive isolation and hybridization. Previous studies have shown that our target species pairs are reproductively isolated, with no viable and fertile hybrids reported to date (Oliveira et al., 1991; Barbadillo & Lapeña, 2003). However, little is known about the intrinsic and extrinsic factors involved in the maintenance of species boundaries in areas where their geographical distributions overlap. Here, we address the role of extrinsic (environmental) factors in

---

**Figure 1** Flowchart of the hypothesis-testing approach followed in this study.
patterns of spatial co-occurrence between Iberian *Alytes* and *Hyla* species, and discuss the implications of our findings in an evolutionary context.

**MATERIALS AND METHODS**

**Study area and data**

The study area was the Iberian Peninsula, the south-western tip of Europe (Fig. 2). With a total area of nearly 600,000 km², it is a physiographically heterogeneous region, comprising the mainland territories of Portugal and Spain, and linked to the European continent by a narrow and mountainous isthmus. It, thus, constitutes a discrete biogeographical unit appropriate for modelling species distributions (Real et al., 2009; Barbosa et al., 2010).

The species distribution data, consisting of presences and absences on 10 × 10 km grid cells with a Universal Transverse Mercator (UTM) projection (Fig. 2), were compiled from Loureiro et al. (2010) for Portugal and from MAGRAMA (2015; including Pleguezuelos et al., 2002) for Spain. Maps of the 10 × 10 km grids and the boundary of the study area were downloaded from the EDIT Geoplatform (Sastre et al., 2009). We used QGIS 2.4 (QGIS Development Team, 2014) for spatial data processing and R 2.11 (R Core Team, 2014) for data management and analysis.

We used 24 potential predictor variables representative of Iberian physiography, climate and human activity (Table 1).
This makes model predictions independent of each species’ presence to absence ratio in the study area, enabling direct model comparison and combination when two or more taxa are analysed (Real et al., 2009; Estrada et al., 2011; Barbosa & Real, 2012).

The selection of variables for each model was made in four phases, all included in the mglm function of the fuzzySim R package (Barbosa, 2015a). First, we pre-selected variables that had a significant bivariate relationship with the target species distribution, accounting for the false discovery rate by correcting the significance of variables under repeated testing (Benjamini & Hochberg, 1995; García, 2003). Then, to avoid multicollinearity and its effects on coefficient estimates (Legendre & Legendre, 2012), among each pair of selected variables that were visibly correlated (Pearson’s \( r > 0.8 \)), we excluded the one that had the weakest relationship with the response, considering both P-value and Akaike’s information criterion (AIC; Akaike, 1973). Using the remaining variables, we built multivariate models with a forward−backward stepwise selection, using AIC to include or exclude variables (Burnham & Anderson, 2002) as is implemented in R. Finally, if there were any variables with non-significant coefficients left in the model (chi-squared test, \( P > 0.05 \)), they were eliminated step by step until all coefficients were significantly different from zero (Crawley, 2007).

We used the fuzzyOverlay function of the fuzzySim package to calculate a fuzzy intersection (Zadeh, 1965) between
the model predictions for the species in each congeneric pair. This intersection represents favourability for the co-occurrence of both species, and has been used previously in comparative distribution modelling (Barbosa & Real, 2010, 2012).

Model analysis and evaluation

We assessed the two different components of performance in each model: discrimination capacity (i.e. the ability of the model to distinguish presences from absences); and calibration or reliability (i.e. the magnitude of the overall deviations of predictions from observed values, and the relationship between predicted probabilities and observed occurrence frequencies). Although the latter component is often neglected, both are equally important in the evaluation of species distribution models (Pearce & Ferrier, 2000; Wintle et al., 2005; Jiménez-Valverde et al., 2013).

The discrimination capacity of the models was measured with several widely used indices, including, for example, the area under the receiver operating characteristic curve (AUC) and the true skill statistic (TSS). The AUC summarizes the discrimination performance of a model over the whole range of prediction thresholds. TSS weights sensitivity (correct classification of presences) and specificity (correct classification of absences) at a given threshold; we chose 0.5 as the threshold for discrimination measures. With favourability models, this equates to using prevalence as a threshold in logistic regression (Real et al., 2006; Acevedo & Real, 2012).

Model reliability was measured with the Hosmer–Lemeshow goodness-of-fit test (Hosmer & Lemeshow, 1980), as well as with $D^2$ (proportion of deviance explained by each model) and the pseudo $R^2$ measures of Cox-Snell, McFadden, Nagelkerke and Tjur. All metrics are implemented in a unified framework in the modEvA R package (Barbosa et al., 2013), including functions AUC, threshMeasures, HLfit, Dsquared and RsqGLM.

Model (environmental niche) comparison

We first analysed the correlations between model predictions, checking for significant similarities or dissimilarities between environmental favourability values for the analysed species and for the co-occurrence of congeneric species. We calculated the significance of these correlations after accounting for spatial autocorrelation with the modified.ttest function of the SpatialPack R package (Osorio et al., 2012), which assesses the correlation between two spatial processes, estimating an effective sample size that takes into account their spatial association (Clifford et al., 1989).

Pairs of models whose predictions were correlated (either positively or negatively) after accounting for spatial autocorrelation were then analysed for niche overlap using Schoener’s $D$ and Warren et al.’s (2008) indices, and for overall similarity based on fuzzy versions (Barbosa, 2015a) of classic binary similarity indices, accounting for shared presence areas (Jaccard, 1901) and for both shared presence and shared absence areas (Baroni-Urbani & Buser, 1976). The latter two indices have associated tables of significant values (Baroni-Urbani & Buser, 1976; Real & Vargas, 1996; Real, 1999). These two indices are traditionally based on categorical presence-absence data, but the fuzzySim function of the fuzzySim package adapts them to work directly with continuous data such as those provided by favourability model predictions, without any loss of quantitative information (Barbosa, 2015a).

RESULTS

Model performance and variables

The four models achieved good overall prediction accuracy, with, for example, AUC values generally above 0.8, which is considered ‘good’, or 0.9, which is considered ‘excellent’ (Swets, 1988); and McFadden’s pseudo-$R^2$ values generally above 0.2, which is considered ‘excellent fit’ (McFadden, 1978) (Fig. 3). The variables included in each model and their coefficient estimates are shown in Table 1. For more detailed statistics on model parameters, see Appendix S1 in Supporting Information.

Favourability for midwife toads was higher in grid cells with higher values of maximum precipitation within a day (MP24) for both species. Relative maximum precipitation (RMP) and the annual air humidity range (HRan) had opposite effects on the favourability for these two species. Favourability for A. obstetricans was also associated with higher values of relative maximum precipitation (RMP) and lower values of potential evapotranspiration (PET), January temperature (TJan) and annual variation in air humidity (HRan), among other variables (Table 1). Favourability of A. cisternasii was associated with lower values of annual precipitation and of how concentrated it is in time (Prec and RMP), and with a smaller annual number of frost days (DFro), in this multivariate context (Table 1; Appendix S1).

Among the tree frogs, the annual number of snow days (DSno), January air humidity (HJan) and annual insolation (Inso) had opposite effects on environmental favourability for the two, while both species were associated with lower maximum daily precipitation (MP24). Favourability for H. molleri was also higher in grid cells with more time-concentrated precipitation (RMP) and lower values of slope (Slop). Favourability of H. meridionalis was associated with lower elevation (AltI), among other variables in the multivariate context of Table 1 (see also Appendix S1). Models were robust to the exclusion of dubious, isolated occurrence records.

Relationships between models

After accounting for spatial autocorrelation, there was a significant negative correlation between the environmental favourability patterns of the analysed midwife toad species.
Figure 3 Discrimination and calibration performance of the environmental favourability models. Measures and plots were obtained with the modEvA R package. AUC: area under the curve. CCR: correct classification rate. TSS (true skill statistic) and kappa were standardized (s) to vary between 0 and 1, so that they can be directly compared with the other measures (Barbosa, 2015b).
distribution range, in the Betic mountains of SE Iberia (Fig. 4). This area has long been occupied by a possible competitor, the morphologically similar, small-range endemic Betic midwife toad *A. dickhilleni* (Pleguezuelos et al., 2002), which was not included in this study because it does not share noticeable co-occurrence areas with the other *Alytes* species. For *A. cisternasii*, we also found favourable conditions outside of its current distribution range in the Iberian south, centre and north-east, where the current absence of this species may reflect dispersal limitations. Favourable areas in the south are separated from this species’ current range by the Guadalquivir river basin (S Spain), which has acted as a major biogeographical barrier for several amphibian taxa (García-Paris et al., 1998; García-Paris & Jockusch, 1999; Carranza & Arnold, 2004; but see Real et al., 2005). On the other hand, genetic data show the existence of several independent glacial refugia for this species in western Portugal during the Pleistocene, from which the present range was colonized through demographic expansions to the north and east, as revealed by a progressive loss of genetic variation along these geographical axes (Gonçalves et al., 2009).

This interpretation probably applies also to *H. molleri*, whose favourable but unoccupied regions in the Iberian south and east probably represent areas to which this species has not yet been able to expand from its western refugia (I. Martínez-Solano, unpublished data). The presence of the congeneric *H. meridionalis* in these areas may represent an additional obstacle to its potential expansion, through interspecific competition. The two Iberian nuclei of *H. meridionalis*, in the south-west and north-east (Fig. 2), originate from two different genetic sources and have probably reached Iberia in different historical times as well (Recuero et al., 2007). In this case, the discrepancies between model predictions and the actual distribution suggest a great expansion potential for the north-eastern nucleus, along the Iberian Mediterranean coast and inland along the Ebro river basin (NE Spain). In central Iberia, however, the Sistema Central mountains seem to present a major dispersal barrier for this species, possibly limiting its expansion (e.g. Sillero, 2009).

**Contrasting co-occurrence patterns in *Alytes* and *Hyla***

Apart from the historical insights obtained from individual species models, the fuzzy intersections of these models (revealing favourability for co-occurrence; Barbosa & Real, 2010, 2012) among each species pair revealed contrasting results that are informative about the mechanisms maintaining species borders in strict sympathy. In midwife toads, we found evidence of niche divergence between *A. obstetricans* and *A. cisternasii*, with very little observed overlap (Fig. 2), despite some environmental potential for co-occurrence as inferred from our models (Fig. 4). These species diverged at least 10 Ma (Gonçalves et al., 2015), and thus it is reasonable to assume that there have been opportunities for secondary contact and hybridization at different times in their
evolutionary history, at least in some parts of their ranges. Thus, given that these two species are currently reproductively isolated (hybridization is extremely rare; Rosa, 1995), and considering their small distributional overlap (Fig. 2, Table 2), it follows that ecological factors may have played a major role in speciation in this group. In anurans, species differences in call characteristics act as a powerful pre-zygotic reproductive isolation mechanism, also leading to assortative mating and thus promoting species divergence (Lemmon, 2009). In Alytes, the calls of the different species are very simple and similar, although playback tests showed differences in female preference between sympatry and allopatry areas, which are consistent with reproductive character displacement under co-occurrence (Márquez & Bosch, 1997). Also, temporal overlap in their breeding seasons is limited (Malkmus, 2004), restricting the opportunities for hybridization where the two species spatially overlap. External factors, here represented by the environmental variables in our models, can also contribute to promote or maintain species differentiation.

In tree frogs, in contrast, we found no evidence of niche divergence and a better match between observed and predicted overlap (Figs 2 & 4), with most favourable areas for the co-occurrence of both species located within the environmental range of H. meridionalis – that is, H. meridionalis, the most recent Iberian resident (Recuero et al., 2007), is occupying a subset of the areas that were also favourable for H. molleri. In this case, assortative mating through audible differences in male calls between the two species acts as an efficient barrier to hybridization (Barbadillo & Lapeña, 2003). Additional post-zygotic barriers, such as reduced fitness or sterility of hybrids, have also been shown to play a role: histological examination of the gonads of hybrid individuals revealed the absence of viable spermatogonia (Barbadillo & Lapeña, 2003). Clearly, environmental factors like those included in our models play a minor role in shaping patterns of Hyla co-occurrence at a finer spatial scale. In fact, in areas where the two species co-occur (for instance, in southern Portugal and along the southern slopes of the Sistema Central mountains in central Spain) it is common to find breeding aggregations of both species in syntopy (e.g. Oliveira et al., 1991; Barbadillo & Lapeña, 2003), with H. meridionalis often occurring in higher abundance (Ferreira, 2006).

We should remark that broad-scale species occurrence patterns are not necessarily matched by patterns on finer, more local scales (Segurado et al., 2012). While the analysed species partially co-occur at broad spatial scales (Fig. 2; Sillero

![Environmental favourability values obtained for the studied species and their fuzzy intersection (favourability for co-occurrence) within each congeneric pair.](image-url)
et al., 2014; IUCN, 2015), microhabitat differences may preclude strict sympatry on a local scale (see Introduction). Finer-scale distribution data from these co-occurrence areas can thus be used, in the future, to analyse these patterns in more detail.

**Evolutionary and conservation implications**

Our study showed how species distribution models can provide relevant information to help reconstruct the evolutionary history of species and of their interactions with other taxa, as well as identify the relative roles of intrinsic versus extrinsic factors in maintaining species boundaries in areas of secondary contact. In addition, from a conservation perspective, our results suggest that conservation efforts for tree frogs may be focused on the areas that are favourable for both species, regarding significant similarity between their favourability patterns; whereas efforts aimed at the conservation of midwife toads will require specific measures tailored separately for each species because of their environmental segregation.

**ACKNOWLEDGEMENTS**

We thank all field workers who provided the data used in this study, and Hugo Rebelo for early discussion of the results. L.R., M.F. and A.M.B. are supported by Fundação para a Ciência e a Tecnologia (FCT), Portugal (contracts SFRH/BPD/93079/2013, SFRH/BD/95202/2013 and IF/00266/2013 respectively; the latter including exploratory project CP1168/C0001). These grants were funded under POPH – QREN – Typology 4.1, by FEDER Funds through the Operational Programme for Competitiveness Factors – COMPETE, and by National Funds (POCI-01-0145-FEDER-006821, unit UID/ BIA/50027). I.M.S. was funded by project ‘Biodiversity, Ecology and Global Change’, co-financed by North Portugal Regional Operational Programme 2007/2013 (ON.2–O Novo Norte), under the National Strategic Reference Framework (NSRF) through the European Regional Development Fund (ERDF), and is currently supported by funding from the Spanish ‘Severo Ochoa’ Program (SEV-2012-0262). C.X. was supported by the National Natural Science Foundation of China (41271197).

**REFERENCES**


SUPPORTING INFORMATION

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

Appendix S1 Model parameters and associated statistics.

BIOSKETCHES

Luis Reino (PhD) and Mário Ferreira (MSc) contributed equally to this work and are considered joint first authors. They are post- and pre-doctoral researchers at CIBIO/InBio, working mainly on the distribution and ecology of birds and amphibians. The team have a common interest in the interactions between climate and biogeographical patterns.

Author contributions: L.R. and P.S. conceived the idea; A.M.B., M.F. and L.R. gathered and analysed the data and produced the results; A.M.B., L.R., I.M.S. and M.F. wrote the paper; I.M.S. and C.X. provided ideas for additional analyses and improved writing and interpretation.

Editor: Krystal Tolley