

# Natural abiotic factors more than anthropogenic perturbation shape the invasion of Eastern Mosquitofish (*Gambusia holbrooki*)

Christina A. Murphy<sup>1,2,4</sup>, Gaël Grenouillet<sup>3,5</sup>, and Emili García-Berthou<sup>1,6</sup>

<sup>1</sup>Institute of Aquatic Ecology, University of Girona, E-17071 Girona, Catalonia, Spain

<sup>2</sup>Department of Fisheries and Wildlife, Oregon State University, Corvallis, Oregon 97331 USA

<sup>3</sup>Université Paul Sabatier, CNRS, ENFA, UMR5174 EDB (Laboratoire Évolution et Diversité Biologique), 118 route de Narbonne, F-31062 Toulouse, France

**Abstract:** Eastern Mosquitofish (*Gambusia holbrooki*) is an invasive and globally widespread species that is considered highly tolerant. We used species distribution models (SDMs) to assess factors, including the role of anthropogenic perturbation, that mediate its invasion on a regional scale. A better understanding of the important large-scale factors may help us identify future areas of concern and potential avenues for control. We built SDMs from presence records and randomly selected pseudo-absences of mosquitofish on a 10- × 10-km grid. We used 10 modeling techniques implemented in the *biomod2* software and ensemble forecasts. Final models contained 12 environmental predictors, including natural environmental factors (elevation, slope, topographic index, precipitation, accumulated flow, temperature mean and range) and anthropogenic perturbation indicators (population density, urban and agricultural land uses, number of local and upstream dams). Elevation, temperature, and accumulated flow most strongly influenced mosquitofish distribution, and mosquitofish were found more frequently in downstream, warmer waters. Anthropogenic features, except the number of upstream dams, were poorer predictors of mosquitofish presence than were natural environmental factors. The best models suggested that mosquitofish are more likely to occur in areas with more dams upstream, but removing this predictor did not strongly affect model results. Restoration efforts or modifications to anthropogenic features appear unlikely to alter mosquitofish distribution patterns, highlighting the importance of preventing introductions to new areas. Mosquitofish have been extensively documented in the Iberian Peninsula, but consensus methods suggest many additional suitable areas from which records were not found. Thus, this highly invasive species is or may become much more widely distributed than current observations in the region.

**Key words:** invasive species, disturbance, distribution modeling, landscape, random forest, Iberian Peninsula

Eastern and Western Mosquitofish (*Gambusia holbrooki* and the closely related *Gambusia affinis*) are among the world's 100 most invasive species (Lowe et al. 2000). They have expanded from their original range in part of the USA and Mexico to become among the most prolific, widespread freshwater fishes in the world (García-Berthou et al. 2005, Pyke 2008). Mosquitofish can survive a wide range of physical conditions and may reproduce very rapidly (Pyke 2008). Their invasions are marked by ecological changes. These 2 poeciliids have been implicated globally in severe amphibian declines (Gamradt and Kats 1996, Baber and Babbitt 2004), local extirpations of native fishes (Alcaraz et al. 2008), and trophic cascade effects (Hurlbert et al. 1972, Hurlbert and Mulla 1981). Predation on zooplankton and insects by mosquitofish has caused cascading changes

to phytoplankton and primary production with effects on light, nutrients, temperature, pH, and dissolved O<sub>2</sub> (Hurlbert et al. 1972, Hurlbert and Mulla 1981). Mosquitofish have been intentionally introduced across the globe, in part because of the frequent misconception that they have a restricted diet and are more efficient predators of mosquitoes than native species (see Pyke 2008 and references therein). Once introduced, we assume that abiotic factors, such as temperature and flow, shaped their ability to become established and to invade a landscape by changing environmental resistance, as predicted by Moyle and Light (1996) based on observations of invasive fishes in California.

Invasibility, the susceptibility of environments to invasion by introduced species, such as mosquitofish, is im-

E-mail addresses: <sup>4</sup>christina.a.murphy@gmail.com; <sup>5</sup>gael.grenouillet@univ-tlse3.fr; <sup>6</sup>emili.garcia@udg.edu

DOI: 10.1086/681948. Received 29 July 2014; Accepted 12 November 2014; Published online 1 May 2015.

Freshwater Science. 2015. 34(3):000–000. © 2015 by The Society for Freshwater Science.

000

portant in understanding current and future species distributions and in explaining the factors that enhance and limit invasions (Lonsdale 1999, Alpert et al. 2000). Temperature, elevation, and water flow strongly affect mosquitofish survival and reproduction at local scales, so we might expect these factors to influence the range and extent of mosquitofish on a landscape scale. For example, temperature is a determinant of the latitudinal extent of mosquitofish through its influence on reproductive rates, overwintering success, and biotic interactions, such as competition with the native Iberian Toothcarp *Aphanius iberus* (Benejam et al. 2009, Carmona-Catot et al. 2013). Higher elevation typically is associated with lower water temperatures, steeper stream habitats, and reduced occurrences of the backwater and side-channel habitats frequently occupied by mosquitofish. The influence of flow on its preferred habitats (Meffe 1984) explains why mosquitofish often are found in areas with low or regulated flow conditions, where warm backwaters can support planktivorous feeding and productivity. We hypothesized that elevation, temperature, and water flow are critical factors in mosquitofish distribution across a landscape.

Human activities, such as land use and construction of dams and reservoirs, and their ecological consequences, can dramatically alter invasibility, but they often are neglected in considerations of invasions over broad extents. This neglect seems counterintuitive because human effects themselves may be easier to quantify over broad extents than their specific ecological consequences. On fine scales, fish assemblages commonly are used in biological indices as metrics of stream health and anthropogenic impacts on stream quality (Hughes et al. 1998), which can influence species distributions (e.g., Vila-Gispert et al. 2002). We expect these relationships to be direct through changes in habitat suitability and indirect through alteration of the biotic community structure and effects on other species.

Mosquitofish can thrive in streams with altered flow regimes and are rare in cold headwater streams (Carmona-Catot et al. 2013), in part because of temperature, but also because they are limnophilic species that do not withstand strong water flows well (Meffe 1984). Mosquitofish are highly tolerant and can thrive in agricultural ditches, sloughs, and urban streams. Thus, these areas may provide additional habitat opportunities (Cabral and Marques 1999, Baber et al. 2002, Northington and Hershey 2006, Simon and Travis 2011). In Mediterranean systems, large dams regulate flow regimes and reduce extreme flow events (Bunn and Arthington 2002). Dams limit high flows and create reservoirs that can provide a constant source of downstream propagules if they are inhabited by mosquitofish. Freshwater exotic fishes often are introduced through reservoirs, from which they spread to rivers (Rahel 2002). We hypothesized that hydrological alterations, such

dams and reservoirs, would favor mosquitofish invasion and that mosquitofish were more likely to be introduced where human activities were more apparent, such as in areas with urban or agricultural development.

We used species distribution models (SDMs) to assess how natural environmental factors and anthropogenic disturbances mediate invasion by mosquitofish, and we used the Iberian Peninsula as a case study. Our objectives were to: 1) understand the factors that mediate mosquitofish invasion, and 2) test the role of anthropogenic disturbances, such as hydrological alteration, in the invasion process. We hypothesized positive relationships between mosquitofish presence and anthropogenically driven hydrological alteration and agricultural land use. As far as we know, we are the first to present an SDM for a mosquitofish species. Mosquitofish have been generally neglected in similar models, except by Lauzeral et al. (2011) who noted some climate niche shifts. Our SDM is one of the first in which anthropogenic landscape features were considered for a freshwater species (see also Markovic et al. 2012, Liang et al. 2013, Maloney et al. 2013). We hope to improve our understanding of factors mediating mosquitofish invasion and to identify areas where mosquitofish may already be present or is likely to invade.

## METHODS

### Data compilation

We obtained 919 unique records from >1000 documented occurrences of mosquitofish in the Iberian Peninsula, an area of ~582,000 km<sup>2</sup> including Spain and Portugal, by reviewing research articles, atlases, and databases (published and unpublished), to build a map with documented presence data (e.g., Doadrio 2002, Ribeiro et al. 2007; see Appendix S1 for additional sources). We considered all mosquitofish (*Gambusia* spp.) records to be the Eastern Mosquitofish (*G. holbrooki*) because recent genetic studies have identified only *G. holbrooki* in the Iberian Peninsula and only *G. holbrooki* is known to have been introduced (Vidal et al. 2010). This distinction is important because grouping of *G. affinis* and *G. holbrooki* as subspecies of *G. affinis* until around 1988 resulted in uncertainty in their respective distributions and inconsistent naming (Wooten et al. 1988, Rauchenberger 1989, Robins et al. 1991), and some references still cite *G. affinis*. We used a resolution of 10 km UTM (Universal Transverse Mercator; i.e., 100 km<sup>2</sup>) because it was the most widely used resolution in the records and could be obtained for all data. Our final grid covered the entire Iberian Peninsula and contained 6138 unique 10 km UTM grid cells. Mosquitofish were recorded in ~15% of the cells (Fig. 1A).

We included variables that have been frequently used in SDMs of freshwater fish, such as elevation, topographic index, flow accumulation, slope, land cover, precipitation,

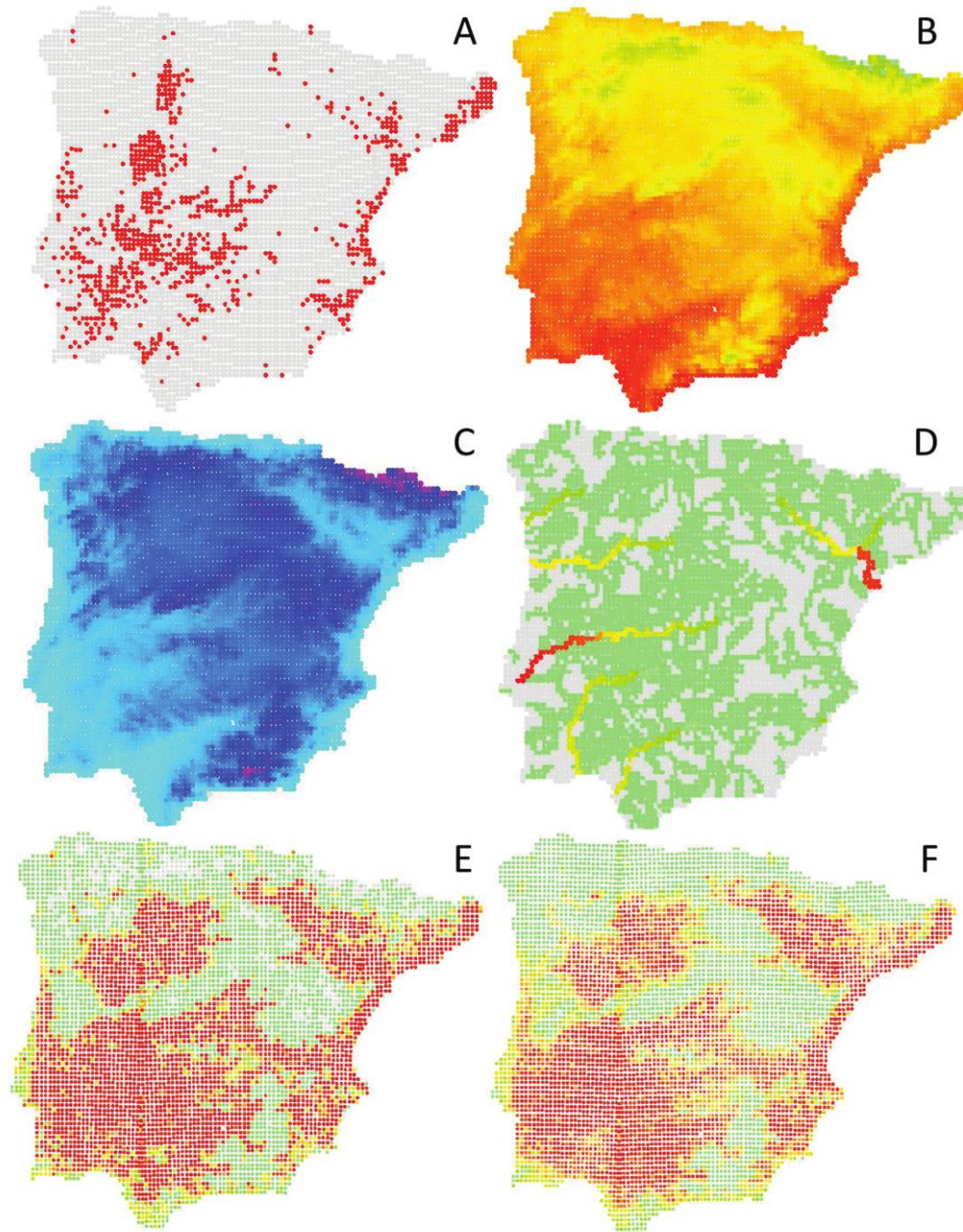


Figure 1. Maps of the Iberian Peninsula showing current records of mosquitofish distribution (A), mean annual temperature (B), elevation (C), total number of upstream dams (D), random forest prediction of mosquitofish probability of presence (E), and ensemble forecast of mosquitofish probability of presence (F). Darker colors are presence, warmer temperatures, higher elevations, more dams, and increased probability of presence, respectively.

annual thermal range (the difference between annual minimum and maximum temperatures), and temperature, as predictors of mosquitofish presence (e.g., McNyset 2005, Chen et al. 2007, DeVaney et al. 2009, Grenouillet et al. 2011). We compiled environmental data layers from on-line databases in ArcGIS (version 10; Environmental Sys-

tems Research Institute, Redlands, California) and did subsequent calculations with its Spatial Analyst toolbox. We considered natural descriptors thought to be related to the establishment, probability, and abundance of mosquitofish (e.g., temperature [Fig. 1B], elevation [Fig. 1C], flow accumulation) and anthropogenic factors (proximity and num-

ber of dams [Fig. 1D], landuse alterations) thought to affect the likelihood of introduction, establishment, or abundance and, thus, likelihood of occupancy. Factors affecting abundance would be expected to influence survival and reproduction and might suggest factors that, at their extremes, are highly suitable or unsuitable for this species.

When environmental data were originally of finer grain than our species data, we used the mean of values within each 10- × 10-km grid cell. As recommended by Dormann et al. (2013), we removed highly correlated variables ( $|r| \geq 0.7$ ), keeping the variable that we expected to contain the most relevant information (e.g., mean annual temperature was retained rather than minimum temperature of the coldest month or maximum temperature of the warmest month when all 3 were correlated). Removing highly correlated variables reduces the resulting variance of regression parameters, improves processing time, reduces errors during processing, and prevents possible misinterpretation of results (Guisan and Thuiller 2005, Elith and Leathwick 2009a). However, if predictor variables have independent ecological relevance, they should be retained. According to a recent, comprehensive review (Dormann et al. 2013), regression-type approaches (e.g., generalized linear models) and machine-learning methods work reasonably well under moderate collinearity and are not outperformed by techniques specifically designed to deal with collinearity (e.g., partial least squares). Therefore, neither elevation nor mean annual temperature were eliminated for the SDMs because, although correlated, both have independent effects on mosquitofish populations (Carmona-Catot et al. 2011). In the Iberian Peninsula, this correlation is presumably exacerbated by topography, including the location of the Pyrenees and Cantabric mountain ranges to the north. Our original environmental matrix had 23 variables, but only 12 were used (Table 1) after elimination of highly correlated predictors.

## Modeling

For fitting species distribution models, we used the BIOMOD computation framework (Thuiller 2003) as implemented in the *biomod2* package (Thuiller et al. 2013) in R (version 3.0.3; R Project for Statistical Computing, Vienna, Austria). We used generalized linear models (GLM), generalized boosting models (also known as boosted regression trees, BRT), classification tree analysis (CTA), mixture discriminant analysis (MDA), random forests (RF), artificial neural networks (ANN), generalized additive models (GAM), surface range envelopes (SRE), multiple adaptive regression splines (MARS), and Maxent (see Olden et al. 2008, Elith and Leathwick 2009b, Franklin and Miller 2009, Dormann et al. 2013 for reviews of these methods). On average, the different modeling methods yield good agreement between observed and predicted distributions, but relative performance of different techniques is idiosyncratic

Table 1. Environmental data compiled for analyses. Thermal range is calculated as the difference between the minimum and maximum annual temperatures. \* indicates layers calculated using ArcGIS from existing data.

Description	Range (minimum–maximum)	Source <sup>a</sup>
Elevation (m)	0–2802	1
Slope (°)	0.0–27.1	1*
Topographic index	–0.6–10.3	1*
Flow accumulation	0.3–129,055.0	1*
Urban land use (%)	0–95	2*
Agriculture land use (%)	0–100	2*
Annual precipitation (mm)	209.1–2402.7	3
Annual mean temperature (°C)	0.6–19.2	3
Thermal range (°C)	5.5–17.0	3*
Population density (individuals/km <sup>2</sup> )	0.3–8749.8	4
Number of local dams	0–19	5, 6*
Total number of dams upstream	0–279	5, 6*

<sup>a</sup> Sources: 1 = Spanish National Center for Geographic Information (<http://centrodedescargas.cnig.es/>); 2 = National Center of Geographical Information (<https://www.cnig.es/>); 3 = Universitat Autònoma de Barcelona, Atlas climático digital de la Península Ibérica (<http://www.opengis.uab.es/>); 4 = DIVA-GIS data (<http://www.diva-gis.org/datadown>); 5 = Melo and Gomes 1992; 6 = Spanish Ministry of Agriculture, Food and Environment (<http://sig.magrama.es/>)

across species (Thuiller 2003). The BIOMOD approach implements several widely used regression techniques (e.g., GLM, GAM, and MARS). It also includes a number of more modern machine learning methods, such as RF and BRT, which offer greater predictive performance for non-linear relationships, compared to the more traditional methods (Olden et al. 2008, Elith and Leathwick 2009b). Combining the results of the different techniques with ensemble forecasting often yields more robust predictions (Araújo and New 2007, Thuiller et al. 2009).

The mosquitofish presence data mostly originated from planned electrofishing quantitative surveys with accurate location records, but we treated the data as presence-only and built pseudo-absences because fish species absence is very difficult to ascertain accurately and probably is subject to a number of potential biases (e.g., contrasting surveying effort; effects of stream size, water conductivity, and flow on detectability; Barbet-Massin et al. 2012). For tests run in *biomod2*, we used 1000 randomly selected pseudo-absences, with no repetition of selection, for 2-run evaluations. These values were estimated based on the areal coverage recommendations provided in Barbet-Massin et al. (2012). We split our data into 70/30 training and test data and randomly selected pseudo-absences. We compared the results to those obtained with repetition and ½ the number

of randomly selected pseudo-absences and noted no obvious deviations. As such, only results with the initial settings are presented below.

We evaluated model accuracy (predictive power) of the different techniques using the following criteria: the area under the receiver operating characteristic (ROC) curve (AUC), Cohen's kappa ( $\kappa$ ), the true skill statistic (TSS), and sensitivity (the proportion of correctly predicted presences) and specificity (the proportion of correctly predicted absences) for each (Allouche et al. 2006, Thuiller et al. 2009), where the binary threshold was defined as the one that maximized the evaluation metric value. As recommended by Allouche et al. (2006), we focused on TSS over  $\kappa$  because the former does not depend on prevalence (the proportion of sites in which the species was recorded as present). We compared the variable contributions from all models to determine the most influential environmental factors. We noted the relative magnitude and order of variable contributions, and we paid special attention to the models with best accuracy. Last, we computed an ensemble forecast built using models with a TSS score greater than 0.7 (RF and Maxent) weighted by their evaluation scores. We also built ensemble forecasts with the same threshold for AUC and  $\kappa$ , only to evaluate these metrics on ensemble performance.

## RESULTS

The model outputs confirmed our hypotheses based on the previous knowledge of mosquitofish ecology and predicted an increasing probability of mosquitofish presence at lower altitudes with lower gradients and higher mean temperatures. The model also predicted a trend toward increasing probability of mosquitofish presence with reduced precipitation and increased thermal range. With the exception of cumulative upstream dams, anthropogenic factors had little to no effect on the probability of mosquitofish presence.

The SDM results, including individual and ensemble models and especially the best-performing model (RF), suggested that mosquitofish could occupy a wider extent than has been documented in the Iberian Peninsula (Fig. 1E, F). The RF model predicted that  $>1/2$  of the Iberian Peninsula, including much of the southwest from where few records exist, appears highly suitable for mosquitofish (Fig. 1E).

The importance of each variable in all 10 model calibrations was more consistent for variables that were either very strong or very weak predictors (Table 2). Elevation was the most influential variable in all models explaining mosquitofish distribution, but its contribution to each model varied. Intermediate variables were not ranked consistently across tests. The 2<sup>nd</sup>-most influential variable varied by test and included mean annual temperature, thermal range, slope, annual precipitation, and flow accumulation. All of those variables, except flow accumulation, appeared

in the top 5 variables by contribution for  $\geq 1/2$  of the tests. Among environmental drivers, topographic index was the poorest predictor and was not notably important in its contribution to any model. At best, it contributed just  $>3\%$ . Flow accumulation also generally contributed little to model building, except in the case of ANN. The other environmental predictors (mean annual temperature, annual precipitation, thermal range, slope, and elevation) were all stronger contributors to  $\geq 1/2$  of the calibrated models (Table 2).

Anthropogenic variables generally were poor predictors compared to abiotic drivers, such as elevation, temperature, and thermal range (Table 2). The cumulative-upstream-dams metric was the only anthropogenic variable ranked in the top 3 predictors (RF, Maxent, and CTA models). Models run without cumulative upstream dams revealed that it had little effect on model performance. Of the land-use measures, agriculture appeared to be a better predictor than urban use. However, % agricultural land use was in the top 5 variables for  $<1/2$  of the tests and always had a low contribution ( $<10\%$ ). Number of local dams and % local urban land use were always poor predictors of mosquitofish presence and were not included in the top 5 variable contributions for any of the 10 tests. Human population density and flow accumulation also were poor predictors, contributing only to ANN and SRE results, which appeared to deviate from the general patterns of variable rankings observed across other tests.

RF and Maxent, followed by BRT, were the best models (based on TSS and AUC criteria; Table 3). RF and Maxent were the only models with TSS scores  $>0.7$ . In addition to their high total scores, RF and Maxent were consistently among the most sensitive and the most specific models judged by all evaluation metrics. In contrast, the ensemble model performed poorly relative to best performing models. Four of the 5 most important predictors were shared by RF, Maxent, and BRT: elevation, slope, total number of upstream dams, and mean annual temperature (Table 2). Thermal range was important in the RF model, whereas annual precipitation was important in the Maxent model. In the RF model, elevation was the strongest contributor (20.7%), followed closely by mean annual temperature (17.5%). Slope (8.3%), total number of upstream dams (7.9%), thermal range (7.3%), flow accumulation (6.2%), and annual precipitation (6%) all contributed similarly, but human population density, land use (urban and agriculture), local dams, and topographic index contributed little.

## DISCUSSION

The Iberian Peninsula presents a unique opportunity to study mosquitofish invasions because it was the site of the first mosquitofish introduction in Europe (in 1921). Mosquitofish quickly invaded the Mediterranean basin after this intentional introduction. By 1948, they appear to

Table 2. Variable importance by modeling technique (GLM = generalized linear models, BRT = boosted regression trees, CTA = classification tree analysis, MDA = mixture discriminant analysis, RF = random forests, ANN = artificial neural networks, GAM = generalized additive models, SRE = surface range envelopes, MARS = multiple adaptive regression splines). The top 5 variables (determined by contribution) for each test are noted with superscripts (a–e).

Modeling technique	Mean annual temperature	Annual precipitation	Flow accumulation	Topographic index	Thermal range	Slope	Elevation	Local dams	Total no. dams upstream	% agricultural land use	% urban land use	Human population density
ANN	1.2	19.9 <sup>c</sup>	20.9 <sup>b</sup>	0.0	9.4	15.9 <sup>d</sup>	45.2 <sup>a</sup>	0.0	1.8	2.6	0.2	10.2 <sup>e</sup>
BRT	6.5 <sup>d</sup>	6.4 <sup>e</sup>	2.0	1.5	7.6 <sup>c</sup>	8.6 <sup>b</sup>	27.0 <sup>a</sup>	0.0	4.7	3.5	0.5	0.5
CTA	42.1 <sup>b</sup>	27.2 <sup>d</sup>	13.8	2.0	13.6	19.1 <sup>e</sup>	52.2 <sup>a</sup>	1.1	28.4 <sup>c</sup>	5.9	5.2	4.8
MDA	10.5 <sup>c</sup>	5.3	0.0	2.7	17.4 <sup>b</sup>	4.9	74.8 <sup>a</sup>	0.0	6.0 <sup>e</sup>	8.2 <sup>d</sup>	0.0	0.0
GAM	6.7 <sup>d</sup>	7.4 <sup>c</sup>	0.0	3.1	15.5 <sup>b</sup>	5.7	81.8 <sup>a</sup>	1.1	2.4	6.6 <sup>e</sup>	0.0	0.0
GLM	11.0 <sup>c</sup>	8.5 <sup>d</sup>	1.0	2.7	17.6 <sup>b</sup>	5.6	88.2 <sup>a</sup>	1.5	2.8	7.3 <sup>e</sup>	0.0	0.0
MARS	27.7 <sup>c</sup>	0.0	0.0	2.1	42.2 <sup>b</sup>	27.1 <sup>d</sup>	62.7 <sup>a</sup>	0.0	5.1	9.5 <sup>e</sup>	0.0	0.0
Maxent	18.5 <sup>b</sup>	15.1 <sup>d</sup>	1.6	0.7	3.9	8.8 <sup>e</sup>	30.8 <sup>a</sup>	0.2	15.2 <sup>c</sup>	3.5	0.6	0.2
RF	17.5 <sup>b</sup>	6.0	6.2	1.6	7.3 <sup>e</sup>	8.3 <sup>c</sup>	20.7 <sup>a</sup>	0.1	7.9 <sup>d</sup>	3.4	0.7	2.4
SRE	2.5	6.1 <sup>b</sup>	3.3 <sup>e</sup>	2.0	2.4	3.6 <sup>d</sup>	8.2 <sup>a</sup>	1.9	1.2	2.7	1.0	5.5 <sup>c</sup>

Table 3. Performance of the different modeling techniques for the species distribution models of Eastern Mosquitofish in the Iberian Peninsula. GLM = generalized linear models, BRT = boosted regression trees, CTA = classification tree analysis, MDA = mixture discriminant analysis, RF = random forests, ANN = artificial neural networks, GAM = generalized additive models, SRE = surface range envelopes, MARS = multiple adaptive regression splines, AUC = area under receiver operating characteristic curve,  $\kappa$  = Cohen's kappa, TSS = true skill statistic, sensitivity = the proportion of correctly predicted presences, specificity = proportion of correctly predicted absences.

Technique	Total score			Sensitivity			Specificity		
	$\kappa$	AUC	TSS	$\kappa$	AUC	TSS	$\kappa$	AUC	TSS
ANN	0.49	0.78	0.41	87.7	71.5	86.2	62.8	70.0	54.4
BRT	0.64	0.92	0.69	92.3	85.1	87.1	71.2	85.2	81.6
CTA	0.49	0.81	0.61	90.0	70.1	83.9	57.6	76.0	77.2
MDA	0.43	0.80	0.44	91.3	73.6	84.4	48.4	73.2	60.0
GAM	0.43	0.82	0.48	90.9	73.7	73.2	50.0	73.6	74.4
GLM	0.40	0.81	0.45	86.2	72.5	74.3	55.2	72.4	70.4
MARS	0.47	0.82	0.49	92.2	75.0	85.1	51.6	74.8	64.4
Maxent	0.77	0.96	0.77	91.5	88.6	91.5	85.8	88.6	85.8
RF	0.88	0.98	0.85	98.9	92.2	93.7	85.2	92.0	91.6
SRE	0.10	NA	0.16	46.1	NA	46.2	70.4	NA	70.0
Ensemble (mean)	0.35	0.83	0.51	65.0	85.6	86.5	80.5	65.9	64.8

have reached a distribution quite similar to their current distribution (see fig. 1 by Benejam et al. 2009). Even in areas where mosquitofish have long been established, their spread has been supplemented by frequent subsequent introductions (Díez-del-Molino et al. 2013). The Iberian Peninsula also has a landscape over which diverse anthropogenic effects can be assessed. In the region, extensive modification and flow regulation have resulted in documented changes to habitat and water-quality characteristics, producing aquatic systems that in some cases are dramatically different from pristine conditions.

Invasibility varies with both natural and anthropogenic disturbances. Freshwater ecosystems are regarded as inherently more susceptible than other ecosystems to invasion (Shea and Chesson 2002), and increases in invasibility probably have resulted from anthropogenic alterations, such as those to water quality, flow regimes, and substrate composition (Mooney and Hobbs 2000, Marchetti and Moyle 2001, Marchetti et al. 2004, Lapointe et al. 2012). The consequences of increased invasibility might be especially important in Mediterranean climates, where extensive human impacts and exotic fish introductions have occurred in regions originally supporting a high proportion of endemic freshwater species (Marr et al. 2010, Ribeiro and Leunda 2012). The Eastern Mosquitofish has been implicated regionally in declines of the endemic, endangered cyprinodontiforms *Valencia hispanica* and *Aphanius* spp. and top-down effects in shallow Mediterranean lakes (Blanco et al. 2004, Caiola and de Sostoa 2005, Alcaraz et al. 2008), reinforcing the importance of understanding its ecology and distribution in conservation and restoration efforts.

Landscape-scale factors that appear to be important to mosquitofish distribution were consistent with our expectations. Mosquitofish prefer low-flow habitats (Meffe 1984); temperature plays a critical role in their survival and reproduction (Benejam et al. 2009, Carmona-Catot et al. 2013) at local scales, and they are excluded from high latitudes (Pyke 2008). Therefore, our finding that elevation was the strongest and most consistent contributor to SDMs seems a natural outcome. We would expect elevation to represent gradients of available side-channel and backwater habitats, temperature, and water velocity and that all of these variables (fewer backwaters, lower temperatures, and water velocity) would make life at high elevations less than ideal for this fish. The consistency across tests with which environmental drivers were more influential than anthropogenic factors leaves us confident of the importance of these broad environmental factors to mosquitofish.

Our results agree with previous reports that machine learning methods such as RF, Maxent, and BRT are among the best-performing (e.g., Elith et al. 2006, Franklin and Miller 2009). The models with these techniques had very good accuracy, sensitivity, and specificity as calculated based on their AUC,  $\kappa$ , and TSS scores. These models strongly support environmental gradients as critical to mosquitofish invasion success. The importance of environmental gradients also is broadly consistent with recent work by Strecker and Olden (2014), who found that phylogenetic clustering in nonnative species was related to natural environmental conditions.

For mosquitofish, human impacts appear to be less important than broad-scale environmental gradients. For example, agricultural land use was not an important pre-

dicator, possibly because the models produce probabilistic maps of spatial occupation without predicting or reflecting mosquitofish densities in occupied areas. Nevertheless, this finding suggests that changing anthropogenic landscape features may yield only limited gains in slowing or preventing the spread of mosquitofish, at least in the Iberian Peninsula. The number of upstream dams is an indicator of flow regulation. In many systems, especially Mediterranean waterways prone to flash flooding, dams dampen the peaks and troughs characteristic of natural flow regimes (Lytle and Poff 2004, Poff et al. 2007). Based on the habitat preferences of mosquitofish, we expected the number of upstream dams to have a stronger effect on models of mosquitofish distribution. It is unlikely that mosquitofish would be able to maintain positions in systems where it would be regularly flushed out of headwaters. In addition, its competitive dominance over local species appears to be lower where salinity fluctuations are more extreme, i.e., in unregulated downstream reaches near estuarine and oceanic conditions (Alcaraz et al. 2008). One could reason that restoring the natural flow regime or removing dams might decrease the abundance and impact of mosquitofish. However, we expect that even if mosquitofish could be temporarily removed from an area, such action would be unlikely to eradicate it, in particular because it could recolonize from refuge habitats, such as ponds and floodplains.

We had reasoned that after its long history of introductions, i.e., a high propagule pressure, mosquitofish may have reached the apex of its invasion in the Iberian Peninsula, and we assumed it was maintaining an extent more or less at equilibrium with abiotic factors. Whether this equilibrium actually exists is unclear because the best-performing models suggest that mosquitofish distribution is grossly underestimated and that its expected distribution is  $>4\times$  what has been recorded. In general, the model results appear to identify tributaries to major river systems and the southwestern corner of the Iberian Peninsula as areas with high habitat suitability, suggesting that these areas might harbor undocumented populations or might already have been invaded by mosquitofish. Its long history of invasion would leave us to suspect that it is already present in at least some of these locations, and we would strongly recommend that researchers and management agencies check these areas for established populations. Mosquitofish are small and may be more easily overlooked than other invasive species. An alternative possibility is that our models are missing a landscape-scale factor that limits mosquitofish distributions in the areas where our current models predict them, leading to an overestimation of suitable area. Based on our ecological knowledge of the species, this possibility seems less likely than a lack of observation or previous introduction, especially because scattered records exist near these areas (Fig. 1D).

Our results suggest that natural abiotic factors are much more important than hydrological disturbance in regulating

mosquitofish presence. Models of climate change predict that the region probably will experience warming, concentrated during the summer months, and markedly decreased precipitation with increased variability (Álvarez Cobelas et al. 2005). Estimates specific to inland waters suggest these changes will have stronger effects on endorheic and groundwater systems, alpine waters, and coastal wetlands than on other water bodies (de Castro et al. 2005, Iglesias et al. 2005). Less predictable flow events have the potential to flush mosquitofish downstream, but we think it unlikely that they would eliminate populations because hydrological regulation (in the form of dams and reservoirs) was less important than expected in our models. However, the potential for warming in uninvaded high-altitude areas is of great concern. Temperature and elevation appear to be extremely important limiting factors for the current distribution of mosquitofish in the region. Warmer temperatures and reduced flows could mean that new areas will become suitable for mosquitofish in the future, thereby aiding its invasion. The relative unimportance of anthropogenic hydrological alterations also suggests that restoration efforts might reduce mosquitofish abundance, but are unlikely to extirpate the species from an area. Thus, prevention of further introductions must become a priority. We urge extreme caution, especially in the areas where mosquitofish presence has been predicted but not observed. Given the importance of abiotic rather than anthropogenic factors as limitations on its distribution, once mosquitofish are introduced in a suitable region, there will be few realistic opportunities to make large changes that can limit its distribution.

## ACKNOWLEDGEMENTS

We gratefully thank everybody who contributed mosquitofish presence data, particularly Ignacio Doadrio, Carlos Fernández-Delgado, Teresa Ferreira, and Filipe Ribeiro. Emily Campbell, Brooke Penaluna, and anonymous referees helped to improve earlier versions of the manuscript. Financial support was provided by the Spanish Ministry of Economy and Competitiveness (projects CGL2009-12877-C02-01, CSD2009-00065, and CGL2013-43822-R), the University of Girona (project SING12/09), and the Government of Catalonia (ref. 2014 SGR 484). The EDB lab was supported by 'Investissement d'Avenir' grants (CEBA, ref. ANR-10-LABX-0025; TULIP, ref. ANR-10-LABX-41).

## LITERATURE CITED

- Alcaraz, C., A. Bisazza, and E. García-Berthou. 2008. Salinity mediates the competitive interactions between invasive mosquitofish and an endangered fish. *Oecologia* 155:205–213.
- Allouche, O., A. Tsoar, and R. Kadmon. 2006. Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology* 43: 1223–1232.
- Alpert, P., E. Bone, and C. Holzapfel. 2000. Invasiveness, invasibility and the role of environmental stress in the spread of



- non-native plants. *Perspectives in Plant Ecology, Evolution and Systematics* 3:52–66.
- Álvarez Cobelas, M., J. Catalán, and D. García de Jalón. 2005. Impacts on inland aquatic ecosystems. Pages 109–142 in J. M. Moreno (editor). A preliminary general assessment of the impacts in Spain due to the effects of climate change. Ministerio de Medio Ambiente, Madrid, Spain. (Available from: [http://www.magrama.gob.es/es/cambio-climatico/temas/impactos-vulnerabilidad-y-adaptacion/full\\_report\\_tcm7-199440.pdf](http://www.magrama.gob.es/es/cambio-climatico/temas/impactos-vulnerabilidad-y-adaptacion/full_report_tcm7-199440.pdf))
- Araújo, M. B., and M. New. 2007. Ensemble forecasting of species distributions. *Trends in Ecology and Evolution* 22:42–47.
- Baber, M. J., and K. J. Babbitt. 2004. Influence of habitat complexity on predator–prey interactions between the fish (*Gambusia holbrooki*) and tadpoles of *Hyla squirella* and *Gastrophryne carolinensis*. *Copeia* 2004:173–177.
- Baber, M. J., D. L. Childers, K. J. Babbitt, and D. H. Anderson. 2002. Controls on fish distribution and abundance in temporary wetlands. *Canadian Journal of Fisheries and Aquatic Sciences* 59:1441–1450.
- Barbet-Massin, M., F. Jiguet, C. H. Albert, and W. Thuiller. 2012. Selecting pseudo-absences for species distribution models: how, where and how many? *Methods in Ecology and Evolution* 3:327–338.
- Benejam, L., C. Alcaraz, P. Sasal, G. Simon-Levert, and E. García-Berthou. 2009. Life history and parasites of the invasive mosquitofish (*Gambusia holbrooki*) along a latitudinal gradient. *Biological Invasions* 11:2265–2277.
- Blanco, S., S. Romo, and M.-J. Villena. 2004. Experimental study on the diet of mosquitofish (*Gambusia holbrooki*) under different ecological conditions in a shallow lake. *International Review of Hydrobiology* 89:250–262.
- Bunn, S. E., and A. H. Arthington. 2002. Basic principles and ecological consequences of altered flow regimes for aquatic biodiversity. *Environmental Management* 30:492–507.
- Cabral, J. A., and J. C. Marques. 1999. Life history, population dynamics and production of eastern mosquitofish, *Gambusia holbrooki* (Pisces, Poeciliidae), in rice fields of the lower Mondego River Valley, western Portugal. *Acta Oecologica* 20:607–620.
- Caiola, N., and A. de Sostoa. 2005. Possible reasons for the decline of two native toothcarps in the Iberian Peninsula: evidence of competition with the introduced Eastern mosquitofish. *Journal of Applied Ichthyology* 21:358–363.
- Carmona-Catot, G., J. Benito, and E. García-Berthou. 2011. Comparing latitudinal and upstream–downstream gradients: life history traits of invasive mosquitofish. *Diversity and Distributions* 17:214–224.
- Carmona-Catot, G., K. Magellan, and E. García-Berthou. 2013. Temperature-specific competition between invasive mosquitofish and an endangered cyprinodontid fish. *PLoS ONE* 8: e54734.
- Chen, P., E. O. Wiley, and K. M. McNyset. 2007. Ecological niche modeling as a predictive tool: silver and bighead carps in North America. *Biological Invasions* 9:43–51.
- de Castro, M., J. Martín-Vide, and S. Alonso. 2005. The climate of Spain: past, present and future scenarios for the 21<sup>st</sup> century. Pages 1–62 in J. M. Moreno (editor). A preliminary general assessment of the impacts in Spain due to the effects of climate change. Ministerio de Medio Ambiente, Madrid, Spain. (Available from: [http://www.magrama.gob.es/es/cambio-climatico/temas/impactos-vulnerabilidad-y-adaptacion/full\\_report\\_tcm7-199440.pdf](http://www.magrama.gob.es/es/cambio-climatico/temas/impactos-vulnerabilidad-y-adaptacion/full_report_tcm7-199440.pdf))
- DeVaney, S. C., K. M. McNyset, J. B. Williams, A. T. Peterson, and E. O. Wiley. 2009. A tale of four “carp”: invasion potential and ecological niche modeling. *PLoS ONE* 4:e5451.
- Díez-del-Molino, D., G. Carmona-Catot, R. M. Araguas, O. Vidal, N. Sanz, E. García-Berthou, and J.-L. García-Marín. 2013. Gene flow and maintenance of genetic diversity in invasive mosquitofish (*Gambusia holbrooki*). *PLoS ONE* 8:e82501.
- Doadrio, I. (editor). 2002. Atlas y Libro Rojo de los Peces Continentales de España. Ministerio de Medio Ambiente, Madrid, Spain.
- Dormann, C. F., J. Elith, S. Bacher, C. Buchmann, G. Carl, G. Carré, J. R. García Marquéz, B. Gruber, B. Lafourcade, P. J. Leitão, T. Münkemüller, C. McClean, P. E. Osborne, B. Reineking, B. Schröder, A. K. Skidmore, D. Zurell, and S. Lautenbach. 2013. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography* 36:27–46.
- Elith, J., C. H. Graham, R. P. Anderson, M. Dudík, S. Ferrier, A. Guisan, R. J. Hijmans, F. Huettmann, J. R. Leathwick, A. Lehmann, J. Li, L. G. Lohmann, B. A. Loiselle, G. Manion, C. Moritz, M. Nakamura, Y. Nakazawa, J. M. M. Overton, A. T. Peterson, S. J. Phillips, K. Richardson, R. Scachetti-Pereira, R. E. Schapire, J. Soberón, S. Williams, M. S. Wisz, and N. E. Zimmermann. 2006. Novel methods improve prediction of species’ distributions from occurrence data. *Ecography* 29: 129–151.
- Elith, J., and J. Leathwick. 2009a. Conservation prioritisation using species distribution modelling. Pages 70–93 in A. Moilanen, K. A. Wilson, and H. P. Possingham (editors). *Spatial conservation prioritization: quantitative methods and computational tools*. Oxford University Press, Oxford, UK.
- Elith, J., and J. R. Leathwick. 2009b. Species distribution models: ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution, and Systematics* 40: 677–697.
- Franklin, J., and J. A. Miller. 2009. *Mapping species distributions: spatial inference and prediction*. Cambridge University Press, Cambridge, UK.
- Gamradt, S. C., and L. B. Kats. 1996. Effect of introduced crayfish and mosquitofish on California newts. *Conservation Biology* 10:1155–1162.
- García-Berthou, E., C. Alcaraz, Q. Pou-Rovira, L. Zamora, G. Coenders, and C. Feo. 2005. Introduction pathways and establishment rates of invasive aquatic species in Europe. *Canadian Journal of Fisheries and Aquatic Sciences* 62:453–463.
- Grenouillet, G., L. Buisson, N. Casajus, and S. Lek. 2011. Ensemble modelling of species distribution: the effects of geographical and environmental ranges. *Ecography* 34:9–17.
- Guisan, A., and W. Thuiller. 2005. Predicting species distribution: offering more than simple habitat models. *Ecology Letters* 8:993–1009.
- Hughes, R. M., P. R. Kaufmann, A. T. Herlihy, T. M. Kincaid, L. Reynolds, and D. P. Larsen. 1998. A process for developing and evaluating indices of fish assemblage integrity. *Canadian Journal of Fisheries and Aquatic Sciences* 55:1618–1631.
- Hurlbert, S., and M. Mulla. 1981. Impacts of mosquitofish (*Gambusia affinis*) predation on plankton communities. *Hydrobiologia* 83:125–151.

- Hurlbert, S. H., J. Zedler, and D. Fairbanks. 1972. Ecosystem alteration by mosquitofish (*Gambusia affinis*) predation. *Science* 175:639–641.
- Iglesias, A., T. Estrela, and F. Gallart. 2005. Impacts on hydric resources. Pages 295–344 in J. M. Moreno (editor). A preliminary general assessment of the impacts in Spain due to the effects of climate change. Ministerio de Medio Ambiente, Madrid, Spain. (Available from: [http://www.magrama.gob.es/es/cambio-climatico/temas/impactos-vulnerabilidad-y-adaptacion/full\\_report\\_tcm7-199440.pdf](http://www.magrama.gob.es/es/cambio-climatico/temas/impactos-vulnerabilidad-y-adaptacion/full_report_tcm7-199440.pdf))
- Lapointe, N., J. Thorson, and P. Angermeier. 2012. Relative roles of natural and anthropogenic drivers of watershed invasibility in riverine ecosystems. *Biological Invasions* 14:1931–1945.
- Lauzeral, C., F. Leprieur, O. Beauchard, Q. Duron, T. Oberdorff, and S. Brosse. 2011. Identifying climatic niche shifts using coarse-grained occurrence data: a test with non-native freshwater fish. *Global Ecology and Biogeography* 20:407–414.
- Liang, L., S. Fei, J. B. Ripy, B. L. Blandford, and T. Grossardt. 2013. Stream habitat modelling for conserving a threatened headwater fish in the Upper Cumberland River, Kentucky. *River Research and Applications* 29:1207–1214.
- Lonsdale, W. M. 1999. Global patterns of plant invasions and the concept of invasibility. *Ecology* 80:1522–1536.
- Lowe, S., M. Browne, S. Boudjelas, and M. De Poorter. 2000. 100 of the world's worst invasive alien species: a selection from the global invasive species database. Invasive Species Specialist Group, Auckland, New Zealand. (Available from: <http://www.issg.org/publications.htm#worst100>)
- Lytle, D. A., and N. L. Poff. 2004. Adaptation to natural flow regimes. *Trends in Ecology and Evolution* 19:94–100.
- Maloney, K. O., D. E. Weller, D. E. Michaelson, and P. J. Cicotto. 2013. Species distribution models of freshwater stream fishes in Maryland and their implications for management. *Environmental Modeling and Assessment* 18:1–12.
- Marchetti, M. P., T. Light, P. B. Moyle, and J. H. Viers. 2004. Fish invasions in California watersheds: testing hypotheses using landscape patterns. *Ecological Applications* 14:1507–1525.
- Marchetti, M. P., and P. B. Moyle. 2001. Effects of flow regime on fish assemblages in a regulated California stream. *Ecological Applications* 11:530–539.
- Markovic, D., J. Freyhof, and C. Wolter. 2012. Where are all the fish: potential of biogeographical maps to project current and future distribution patterns of freshwater species. *PLoS ONE* 7:e40530.
- Marr, S. M., M. P. Marchetti, J. D. Olden, E. García-Berthou, D. L. Morgan, I. Arismendi, J. A. Day, C. L. Griffiths, and P. H. Skelton. 2010. Freshwater fish introductions in mediterranean-climate regions: are there commonalities in the conservation problem? *Diversity and Distributions* 16:606–619.
- McNyset, K. M. 2005. Use of ecological niche modelling to predict distributions of freshwater fish species in Kansas. *Ecology of Freshwater Fish* 14:243–255.
- Meffe, G. K. 1984. Effects of abiotic disturbance on coexistence of predator-prey fish species. *Ecology* 65:1525–1534.
- Melo, F. G., and A. S. Gomes (editors). 1992. Large dams in Portugal. Portuguese National Committee on Large Dams, Lisbon, Portugal. (Available from: <http://www.worldcat.org/title/large-dams-in-portugal/oclc/246906036>)
- Mooney, H. A., and R. J. Hobbs. 2000. *Invasive species in a changing world*. Island Press, Washington, DC.
- Moyle, P. B., and T. Light. 1996. Fish invasions in California: do abiotic factors determine success? *Ecology* 77:1666–1670.
- Northington, R. M., and A. E. Hershey. 2006. Effects of stream restoration and wastewater treatment plant effluent on fish communities in urban streams. *Freshwater Biology* 51:1959–1973.
- Olden, J. D., J. J. Lawler, and N. L. Poff. 2008. Machine learning methods without tears: a primer for ecologists. *Quarterly Review of Biology* 83:171–193.
- Poff, N. L., J. D. Olden, D. M. Merritt, and D. M. Pepin. 2007. Homogenization of regional river dynamics by dams and global biodiversity implications. *Proceedings of the National Academy of Sciences of the United States of America* 104:5732–5737.
- Pyke, G. H. 2008. Plague minnow or mosquito fish? A review of the biology and impacts of introduced *Gambusia species*. *Annual Review of Ecology, Evolution, and Systematics* 39:171–191.
- Rahel, F. J. 2002. Homogenization of freshwater faunas. *Annual Review of Ecology and Systematics* 33:291–315.
- Rauchenberger, M. 1989. Systematics and biogeography of the genus *Gambusia* (Cyprinodontiformes, Poeciliidae). *American Museum Novitates* 2951:1–74.
- Ribeiro, F., R. Beldade, M. Dix, and J. Bochechas. 2007. Carta Piscícola Nacional. Direcção Geral dos Recursos Florestais-Fluviatilis, Lisboa, Portugal (in Portuguese). (Available from: <http://www.cartapiscicola.org>)
- Ribeiro, F., and P. M. Leunda. 2012. Non-native fish impacts on Mediterranean freshwater ecosystems: current knowledge and research needs. *Fisheries Management and Ecology* 19:142–156.
- Robins, C. R., R. M. Bailey, C. E. Bond, J. R. Brooker, E. A. Lachner, R. N. Lea, and W. B. Scott. 1991. *Common and scientific names of fishes from the United States and Canada*. American Fisheries Society, Bethesda, Maryland.
- Shea, K., and P. Chesson. 2002. Community ecology theory as a framework for biological invasions. *Trends in Ecology and Evolution* 17:170–176.
- Simon, T. N., and J. Travis. 2011. The contribution of man-made ditches to the regional stream biodiversity of the New River watershed in the Florida panhandle. *Hydrobiologia* 661:163–177.
- Strecker, A. L., and J. D. Olden. 2014. Fish species introductions provide novel insights into the patterns and drivers of phylogenetic structure in freshwaters. *Proceedings of the Royal Society of London Series B: Biological Sciences* 281:20133003.
- Thuiller, W. 2003. BIOMOD—optimizing predictions of species distributions and projecting potential future shifts under global change. *Global Change Biology* 9:1353–1362.
- Thuiller, W., D. Georges, and R. Engler. 2013. biomod2: ensemble platform for species distribution modeling. R package version 3.1-25. R Project for Statistical Computing, Vienna, Austria.
- Thuiller, W., B. Lafourcade, R. Engler, and M. B. Araújo. 2009. BIOMOD—a platform for ensemble forecasting of species distributions. *Ecography* 32:369–373.
- Vidal, O., E. García-Berthou, P. Tedesco, and J.-L. García-Marín. 2010. Origin and genetic diversity of mosquitofish (*Gambusia holbrooki*) introduced to Europe. *Biological Invasions* 12:841–851.
- Vila-Gispert, A., García-Berthou, E., and R. Moreno-Amich. 2002. Fish zonation in a Mediterranean stream: effects of human disturbances. *Aquatic Sciences* 64:163–170.
- Wooten, M. C., K. T. Scribner, and M. H. Smith. 1988. Genetic variability and systematics of *Gambusia* in the southeastern United States. *Copeia* 1988:283–289.