Received: 27 May 2019 | Revised: 8 March 2020 | Accepted: 25 March 2020

DOI: 10.1111/fwb.13513

ORIGINAL ARTICLE

Freshwater Biology WILEY

Historical, human, and environmental drivers of genetic diversity in the red swamp crayfish (*Procambarus clarkii*) invading the Iberian Peninsula

Lucía Acevedo-Limón[1] | Francisco J. Oficialdegui[2] | Marta I. Sánchez[3] | Miguel Clavero[1]

[1]Department of Conservation Biology, Doñana Biological Station (EBD-CSIC), Seville, Spain

[2]Department of Wetland Ecology, Doñana Biological Station (EBD-CSIC), Seville, Spain

[3]Department of Plant Biology and Ecology, Faculty of Biology, University of Seville, Seville, Spain

Correspondence
Lucía Acevedo-Limón, Estación Biológica de Doñana–CSIC, Avenida Américo Vespucio 26, 41092 Sevilla, Spain.
Email: luciaacevedo294@gmail.com

Funding information
Andalusian Government, Grant/Award Number: RNM-936; PICS, Grant/Award Number: PIC2015FR4 and PICS07360

Abstract

1. Patterns of genetic diversity in invasive populations can be modulated by a range of factors acting at different stages of the invasion process, including the genetic composition of the source population(s), the introduction history (e.g. propagule pressure), the environmental suitability of recipient areas, and the features of secondary introductions.
2. The North American red swamp crayfish, *Procambarus clarkii*, is one of the most widely introduced freshwater species worldwide. It was legally introduced into Spain twice, near the city of Badajoz in 1973 and in the Guadalquivir marshes in 1974. Thereafter the species rapidly colonised almost the entire Iberian Peninsula.
3. We used seven nuclear microsatellites to describe the genetic diversity and structure of 28 locations distributed across the Iberian Peninsula and to explain the expansion process of the red swamp crayfish. Additionally, we analysed the relationship between environmental suitability and genetic diversity of the studied locations.
4. The red swamp crayfish had a clear spatial genetic structure in the Iberian Peninsula, probably determined by the two independent introduction events in the 1970s, which produced two main clusters separated spatially, one of which was dominant in Portugal and the other in Spain.
5. The human-mediated dispersal process seemed to have involved invasion hubs, hosting highly genetically diverse areas and acting as sources for subsequent introductions. Genetic diversity also tended to be higher in more suitable environments across the Iberian Peninsula.
6. Our results showed that the complex and human-mediated expansion of the red swamp crayfish in the Iberian Peninsula has involved several long- and short-distance movements and that both ecological and anthropogenic factors have shaped the genetic diversity patterns resulting from this invasion process. Early detection of potential invasion hubs may help to halt multiple short-distance translocations and thus the rapid expansion of highly prolific invasive species over non-native areas.

Lucía Acevedo-Limón and Francisco J. Oficialdegui contributed equally to this work.

KEYWORDS
environmental suitability, genetic structure, human-mediated dispersal, invasive species, microsatellite, multiple introduction

1 | INTRODUCTION

Biological invasions are one of the main threats to biodiversity globally (Bellard, Cassey, & Blackburn, 2016). The intensification of global trade and human movements, as well as the increase of activities such as aquaculture, pet trade, or gardening, have led to an acceleration of the global-scale exchange of biota (Hulme et al., 2008; Ricciardi, 2007), which is blurring the traditionally described biogeographical barriers (Capinha, Essl, Seebens, Moser, & Pereira, 2015). The number of species introduced outside their native ranges has been increasing in recent decades and is expected to keep growing (Seebens et al., 2017). Only a fraction of the introduced species is able to establish self-sustained populations, thrive, and spread, and only a fraction among them cause biodiversity losses, disruptions of ecosystem functioning, and economic impacts (Walsh, Carpenter, & Vander Zanden, 2016). Understanding why some introduced species succeed and become invasive, while other fail, is a central topic in invasion science (Blackburn & Duncan, 2001; Blackburn, Prowse, Lockwood, & Cassey, 2013; Facon et al., 2006).

The genetic diversity of introduced populations can influence their ability to adapt to novel environments and, thus, determine their invasiveness (Lavergne & Molofsky, 2007; but see Bossdorf, Richards, & Pigliucci, 2008; Hawes et al., 2018). Biological invasions are a multistep process often described as a series of stages (transport, introduction, establishment and spread) separated by different barriers that can impede the progress of an invasion (Blackburn et al., 2011). Overcoming each of these barriers can generate population bottlenecks and alter the genetic diversity patterns in invasive populations (Hardesty et al., 2012; Okada, Lyle, & Jasieniuk, 2009). Different factors can modulate the intensity of population bottlenecks in each barrier of the invasion process, including genetic diversity of the source population, propagule pressure, environmental suitability of the recipient area, and/or the characteristics of secondary introductions. Genetic admixture (hereafter admixture) occurs when multiple divergent genetic lineages come into contact and interbreed, increasing the genetic diversity of a population, as can occur in the source population of the native range before the transport stage (van Boheemen et al., 2017; Dlugosch & Parker, 2008; Oficialdegui et al., 2019; Rius & Darling, 2014). During the introduction stage, propagule pressure (i.e. number of introduction events, inoculum size or both) modulates resulting genetic diversity patterns since more introduction events and/or a large number of introduced individuals promote higher genetic diversity in the introduced population (Blackburn et al., 2013; Drolet & Locke, 2016). During the establishment stage, biotic (i.e. niche competition) and abiotic (i.e. environmental suitability) factors can affect the genetic diversity of an introduced population through modulation of survival and its associated population bottleneck (Banks et al., 2013; Ellegren & Galtier, 2016). As such, the environmental suitability refers to the climatic and physiographic variables of the introduced range. During the spread stage, founding events, involving all the previous cited modulators of genetic diversity, take place whenever a secondary introduction occurs (i.e. the source population being itself introduced). Therefore, range expansions are generally associated with decreasing genetic diversity (i.e. allelic richness and expected heterozygosity) along the expansion front (Austerlitz, Jung-muller, Godelle, & Gouyon, 1997; Excoffier, Foll, & Petit, 2009).

The red swamp crayfish (*Procambarus clarkii*), native to north-eastern Mexico and south-central U.S.A., has been broadly introduced around the world, to the point that it is present in up to 40 countries of four continents (Oficialdegui, Sánchez, & Clavero, 2020). It was intentionally introduced to southern Spain in the early 1970s, through two independent shipments from Louisiana (U.S.A.). Both introductions had legal authorisation and were motivated by the high socioeconomic value that crayfish was attaining in Spain (Clavero, 2016). The first introduction took place in Badajoz (Spain) in 1973, and involved the release of around 300 individuals, the survivors of an original batch of 500 crayfish (Habsburgo-Lorena, 1978). One year later, a larger batch (around 500 kg) was imported to the marsh area of the Lower Guadalquivir River (Puebla del Río, Seville), although only 100 kg (around 6,500 individuals) survived. The red swamp crayfish immediately established self-sustained and abundant populations in the initial introduction areas and rapidly spread over the Iberian Peninsula (Gutiérrez-Yurrita et al., 1999; Oficialdegui et al., 2020), aided by both intrinsic traits (e.g. short life cycle, high fecundity, high environmental tolerance; Geiger, Alcorlo, Baltanas, & Montes, 2005) and by multiple (arguably, thousands) and uncontrolled secondary introductions (Clavero, 2016; Oficialdegui et al., 2019). Shortly after the introduction, by 1982, there were already reports of the red swamp crayfish in the Tablas de Daimiel National Park and the Ebro Delta (some 320 and 730 km straight-line, respectively, from the Lower Guadalquivir introduction site) (Clavero, 2016; Oficialdegui et al., 2020). Once introduced and established, this species often becomes dominant in the occupied freshwater habitats, producing severe ecological impacts and losses of ecosystem services (Gherardi, 2006; Souty-Grosset et al., 2016).

In this study, we examine the spatial patterns of genetic diversity of the red swamp crayfish in the Iberian Peninsula to test various hypotheses about the invasion history. Based on the analysis of nuclear microsatellites, we aim to analyse the present-day genetic structure of the red swamp crayfish, to then explore the drivers that may have modulated the dynamics of genetic diversity during the invasion process. The main questions we address are: (1) is there a relationship between the number and size of initial introduction events and the present-day genetic diversity?; (2) how was the pattern of spread of the swamp crayfish among the Iberian Peninsula?; and (3) is there a relationship between environmental suitability and genetic diversity?

TABLE 1 Information on the 28 locations of red swamp crayfish (*Procambarus clarkii*) surveyed in the Iberian Peninsula, including name, code (as in Table 2 and Figure 1), number of collected individuals (*N*), geographical coordinates, type of habitat, environmental suitability (as reported by Capinha & Anastácio, 2011), straight-line distance (km) to the original introduction focus (step distance) and minimum distance between those to the introduction focus or to the nearest invasion hub (hub distance; see Figure 2b)

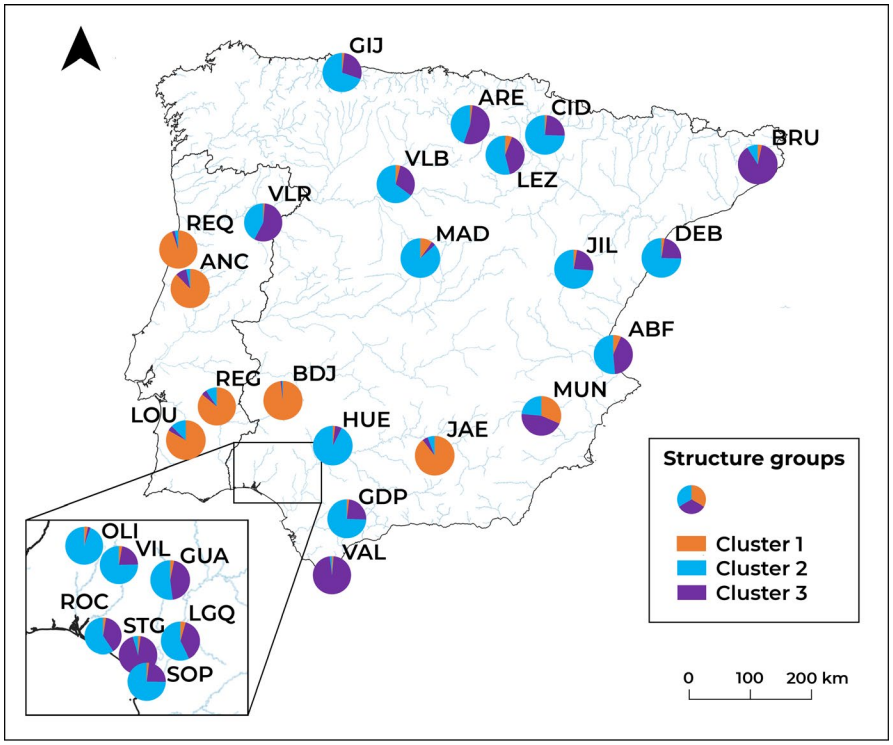
Locations	Code	*N*	Lat	Lon	Habitat	Enviromental suitability	Step dist	Hub dist
Albufera	ABF	26	39.184	−0.192	Rice Field	0.873	566	0
Ança	ANC	30	40.160	−8.306	River	0.803	206	206
Arreo	ARE	20	42.779	−2.991	Lake	0.764	683	382
Brugent	BRU	30	42.021	2.362	River	0.808	927	213
Cidacos	CID	29	42.274	−1.373	River	0.825	708	274
Delta del Ebro	DEB	28	40.783	0.690	Rice Field	0.891	718	0
Badajoz	BDJ	31	38.899	−6.871	Ponds	0.833	0	0
Guadalporcún	GDP	50	36.565	−5.213	River	0.731	75	75
Gijón	GIJ	13	43.321	−5.382	Pond	0.886	713	611
Guadiamar	GUA	50	37.392	−6.134	River	0.801	59	59
Hueznar	HUE	40	37.556	−5.415	River	0.767	98	98
Jaén	JAE	30	37.494	−3.441	River	0.836	301	301
Jiloca	JIL	15	40.544	−1.293	River	0.538	568	171
Leza	LEZ	30	42.263	−2.184	Stream	0.790	676	316
Lower Guadalquivir	LGQ	49	37.755	−6.959	Rice Field	0.894	0	0
Lousal	LOU	30	38.014	−8.255	River	0.857	164	164
Madrid	MAD	30	40.400	−4.056	Pond	0.646	435	350
Mundo	MUN	21	38.273	−1.462	Stream	0.678	415	157
Olivargas	OLI	50	37.471	−6.486	River	0.826	93	93
Reguengos	REG	30	38.284	−7.312	River	0.836	71	71
Requeixo	REQ	30	40.353	−8.313	River	0.786	233	233
Rocina	ROC	30	37.101	−6.372	Stream	0.826	41	41
Sopetón	SOP	20	36.573	−6.266	Lagoon	0.874	32	32
Sotogrande	STG	50	37.097	−6.463	Lake	0.883	27	27
Valle	VAL	50	36.050	−5.414	River	0.737	124	124
Villar	VIL	32	37.412	−6.433	River	0.825	79	79
Valoria la buena	VLB	29	41.801	−4.588	Stream	0.729	536	454
Vila-Rica	VLR	30	41.135	−7.055	Stream	0.744	462	462
Total		903						

We hypothesise that populations originated from Lower Guadalquivir would have a higher genetic diversity than the ones originated from Badajoz, as the inoculum size was around 20 times larger in the Lower Guadalquivir. Besides, genetic patterns within the Iberian Peninsula would be mainly explained by human-mediated dispersal, with a negligible influence of natural dispersal. We also consider two dispersion processes: the jump-dispersal and the invasion hub scenarios. The jump-dispersal scenario assumes that the spread has occurred through successive small-scale secondary introductions, supposing that genetic diversity would tend to diminish with increasing distances to the initial foci, due to the accumulation of genetic bottlenecks at each secondary introduction. The invasion hub scenario involves also large-scale translocations and relevant sources other than the initial foci (i.e. the invasion hubs). It supposes that the high genetic diversity in the invasion hubs could enhance genetic diversity in neighbouring introduced populations. Finally, we hypothesise that suitable environmental conditions would reduce the intensity of population bottlenecks, so that crayfish introduced in suitable areas would present higher genetic diversity than those in unsuitable areas.

2 | METHODS

2.1 | Sample collection, DNA extraction, and microsatellite genotyping

A total of 903 adult red swamp crayfish were collected from 28 locations distributed across the Iberian Peninsula (Table 1; Figure 1). A



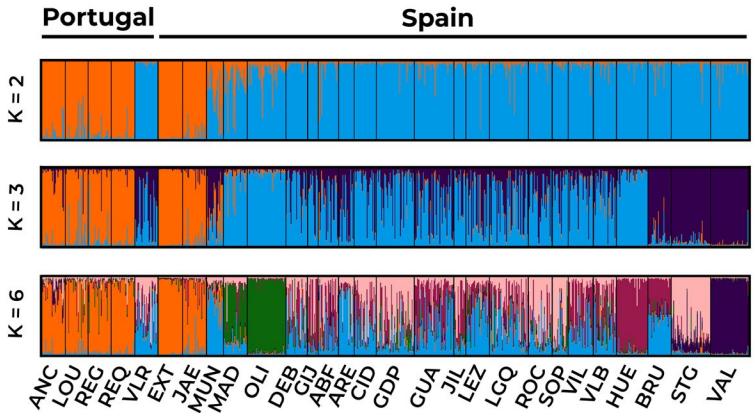


FIGURE 1 Genetic structure of the red swamp crayfish in the Iberian Peninsula, as resulting from STRUCTURE outputs. The upper map shows the spatial distribution of the 28 locations and the proportion of association to each of the genetic clusters defined for the most plausible *K* value (*K* = 3). Lower panels show the probability of assignment of red swamp crayfish individuals to the genetic clusters defined for plausible *K* values (*K* = 2, *K* = 3 and *K* = 6, after the Δ*K* method, Figure S2). In these panels, each vertical line represents an individual, with individuals being grouped by locations (codes as in Table 1), and genetic clusters are represented by different colours

piece of abdominal muscle tissue was extracted from each crayfish and stored in 96% ethanol at room temperature until subsequent analyses.

Total genomic DNA was extracted from approximately 10 mg of dried muscle tissue using a modified DNA salt-extraction protocol (Aljanabi, 1997) containing NaCl 25 mM, Tris 12.5 mM (pH 8.0), EDTA 12.5 mM (pH 8.0), 31.5 µl SDS 10%, 230 µl deionised water, and proteinase K. After overnight incubation at 34°C, DNA samples were extracted with a Tecan robot, Freedom Evo model. Resulting DNA was diluted 1:10 and preserved at −20°C for genotyping analyses. We designed two multiplex polymerase chain reactions (PCRs) for fragment analysis, with Mix 1 (PCSH0002, PCSH0006, PclG-17, PclG-29) and Mix 2 (PCSH0038, PCSH0065, PclG-15, PclG-48) containing microsatellite loci previously developed by Belfiore and May (2000) and Jiang et al. (2015). A multiplex PCR was performed on both Mix 1 and Mix 2 (Table S1). All PCR amplifications were performed in 15-µl reactions containing 4 µl of template DNA, 3 µl buffer 5× PROMEGA, 2.5 mM dNTP, 25 mM $MgCl_2$, 2 µl of primer mix (forward primer end-labelled with [^{32}P] γ ATP), 0.75 U *Taq* polymerase PROMEGA, and deionised water up to the final volume of 15 µl. The thermocycling regime of the Mix 1 consisted of an initial denaturation step at 95°C for 3 min, followed by eight cycles of denaturing at 95°C for 30 s, annealing at 60°C (decreasing 1°C for each cycle) for 30 s, and extension at 72°C for 30 s, followed by 23 cycles of 95°C for 30 s, 52°C for 30 s and 72°C for 30 s with a final extension at 72°C for 10 min. Thermocycling conditions of the Mix 2 were 95°C for 3 min followed by 10 cycles of 95°C for 30 s, 60°C (decreasing 1°C for each cycle) for 30 s, 72°C for 30 s, followed by 23 cycles of 95°C for 30 s, 50°C for 30 s, 72°C for 30 s with a final extension at 72°C for 10 min. Genotyping of amplified products was performed by using an ABI3130xl Genetic Analyser (Applied Biosystem, UK) and allele size was determined using the Genescan 500-LIZ size standard and electrophoretograms were scored in Genemapper version 4.0

(Applied Biosystems). All peaks were manually verified by the lead author to ensure genotyping accuracy.

2.2 | Genetic structure and diversity

MICROCHECKER v.2.2.3 was used to assess the presence of null alleles, large allele drop-outs and scoring errors due to stuttering (van Oosterhout, Hutchinson, Wills, & Shipley, 2004). GENEPOP v.4.7.0 software (Rousset, 2008) was used to detect deviation from Hardy–Weinberg equilibrium (HWE) and linkage disequilibrium between pairs of loci and each locus across locations. While HWE test provided possible departures from equilibrium in our locations, which may indicate systematic genotyping errors and other biases (Salanti, Amountza, Ntzani, & Joannidis, 2005); linkage disequilibrium test was used to assess the independence between analysed loci. Exact tests were used with specified Markov chain parameters of 10,000 dememorisation steps, followed by 5,000 batches of 5,000 iterations per batch. Statistical significance levels were adjusted according to Bonferroni's procedure to counteract the problem of multiple testing in HWE and linkage disequilibrium (Rice, 1989).

To characterise the genetic diversity of the red swamp crayfish in the Iberian Peninsula, we estimated the total number of observed alleles (N_A), the effective number of alleles (N_E), the expected and the observed heterozygosity (H_E and H_O, respectively) and the inbreeding coefficients (F_{IS}) for each locus in each location by using GENALEX v.6.503 software (Peakall & Smouse, 2012). The allelic richness (A_R) and the number of private alleles (P_A) were calculated with ADZE software (Szpiech, Jakobsson, & Rosenberg, 2008), a rarefaction method to be able to compare locations with different sampling sizes. To infer the genetic differentiation among locations, pairwise F_{ST} values were calculated by using ARLEQUIN v.3.1 (Excoffier, Laval, & Schneider, 2005). Bonferroni's correction was performed to adjust the significance for multiple pairwise comparisons in F_{ST} values (Rice, 1989).

BOTTLENECK v.1.2.02 was used to identify locations that have recently experienced a significant reduction in effective population size (Piry, Luikart, & Cornuert, 1999). This software performs a test of heterozygosity based on the assumption that the number of alleles decreases faster than the heterozygosity when a population experience a bottleneck. The stepwise-mutation (SMM) and two-phased (TPM) models with 10,000 replicates were used to test population bottlenecks. Variance for TPM was set to 30 and the proportion of SMM in TPM was set to 80%. The Wilcoxon's test was used to establish whether the number of loci showing heterozygosity excess was significantly greater than expected in locations at equilibrium.

Isolation by distance analysis was used to evaluate the relationship between genetic (F_{ST}) and geographic (based on X-Y coordinates) distances among pairs of locations (Wright, 1943). A Mantel test with 100,000 replicates was performed using *ade4* package in R software (Dray & Dufour, 2007). To calculate the geographic distances among Iberian locations we used the *geosphere* (Hijmans, Williams, & Vennes, 2017) and *Imap* (Wallace, 2015) packages in R v3.2.3 (R Development Core Team, 2014).

STRUCTURE v.2.3.4 was used to characterise the genetic structure of red swamp crayfish in the Iberian Peninsula, and particularly to test whether the two introduction foci can explain the present-day observed genetic structure (Pritchard, Stephens, &

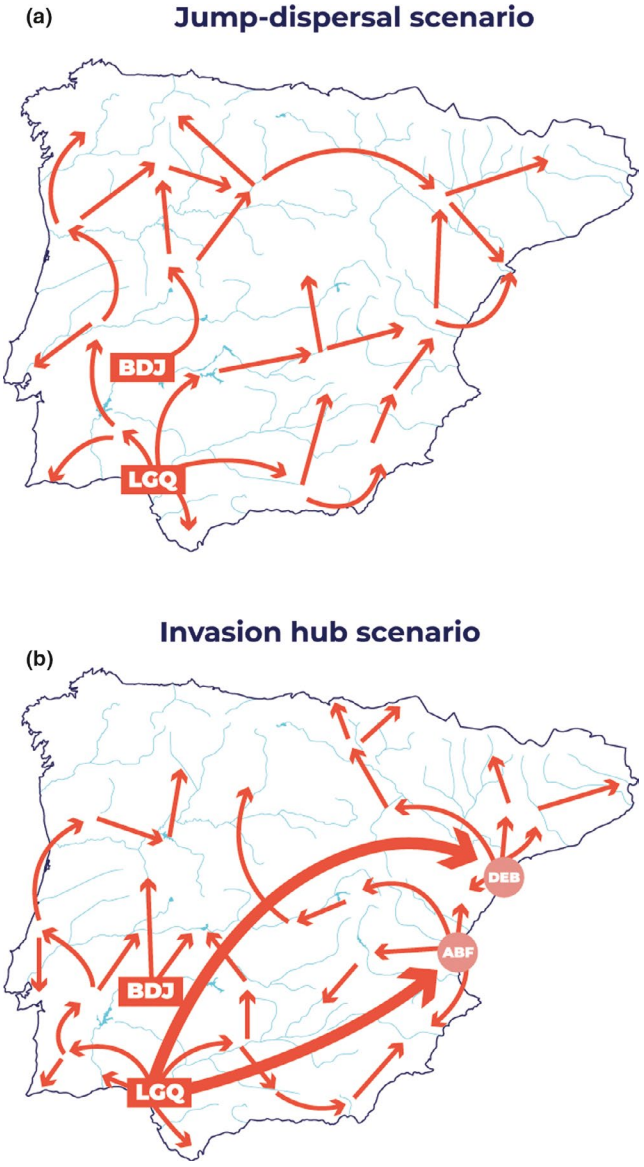


FIGURE 2 Schematic representation of plausible dispersal patterns of the red swamp crayfish (*Procambarus clarkii*) across the Iberian Peninsula. In the jump-dispersal scenario (a), the accumulation of bottlenecks due to successive introduction events would involve a reduction of genetic diversity with increasing distance from the introduction focus (Badajoz, Lower Guadalquivir). Contrastingly, in the invasion hub scenario (b), long-distance transport of genetically diverse crayfish batches (e.g. due to high propagule pressure, which is represented by arrow thickness) could have generated invasion hubs (Valencia Albufera, ABF; Ebro Delta, DEB), acting as additional sources for secondary introductions. Thus, genetic diversity would decrease with increasing distances to either original introduction foci or to invasion hubs.

TABLE 2 Summary of genetic diversity values for seven microsatellite loci in the 28 sampled locations of red swamp crayfish distributed across the Iberian Peninsula

Sampled locations		*N*	N_A	N_E	P_A	A_R	H_O	H_E	F_{IS}
ABF	Mean	25.429	8.571	4.947	0.261	7.138	0.736	0.728	−0.022
	±*SE*	0.149	0.644	0.592	0.137	1.112	0.022	0.026	0.024
ANC	Mean	30.000	6.143	3.606	0.103	5.364	0.605	0.687	0.114
	±*SE*	0.000	0.400	0.294	0.091	0.684	0.032	0.020	0.044
ARE	Mean	19.857	4.857	2.632	0.000	4.472	0.568	0.590	0.015
	±*SE*	0.071	0.442	0.146	0.000	0.710	0.016	0.023	0.033
BDJ	Mean	31.000	5.286	2.497	0.000	4.444	0.465	0.560	0.219
	±*SE*	0.000	0.340	0.152	0.000	0.466	0.051	0.027	0.060
BRU	Mean	30.000	4.857	3.280	0.071	4.554	0.619	0.624	0.001
	±*SE*	0.000	0.335	0.355	0.063	0.647	0.028	0.030	0.015
CID	Mean	29.000	7.571	4.682	0.013	6.347	0.768	0.749	−0.028
	±*SE*	0.000	0.697	0.388	0.012	0.980	0.021	0.019	0.017
DEB	Mean	28.000	8.714	5.046	0.037	7.184	0.760	0.783	0.034
	±*SE*	0.000	0.531	0.340	0.035	0.863	0.025	0.012	0.021
GDP	Mean	49.857	7.857	4.263	0.127	6.062	0.714	0.714	−0.002
	±*SE*	0.071	0.694	0.455	0.123	0.911	0.021	0.022	0.012
GIJ	Mean	12.571	6.000	3.726	0.000	5.921	0.758	0.709	−0.066
	±*SE*	0.101	0.267	0.202	0.000	0.525	0.029	0.018	0.025
GUA	Mean	49.857	9.429	4.753	0.045	7.045	0.667	0.764	0.120
	±*SE*	0.071	0.625	0.357	0.034	0.813	0.012	0.015	0.020
HUE	Mean	39.571	7.143	4.139	0.046	5.978	0.741	0.750	0.016
	±*SE*	0.149	0.369	0.151	0.046	0.500	0.019	0.010	0.014
JAE	Mean	30.000	5.143	2.712	0.059	4.381	0.505	0.603	0.161
	±*SE*	0.000	0.335	0.140	0.059	0.511	0.027	0.023	0.036
JIL	Mean	15.000	6.571	4.781	0.000	6.416	0.724	0.759	0.045
	±*SE*	0.000	0.434	0.364	0.000	0.842	0.029	0.019	0.030
LEZ	Mean	30.000	8.286	4.589	0.027	6.846	0.714	0.755	0.052
	±*SE*	0.000	0.508	0.384	0.026	0.762	0.016	0.014	0.016
LGQ	Mean	47.286	10.571	6.079	0.126	7.932	0.787	0.795	0.009
	±*SE*	0.237	0.770	0.600	0.089	1.043	0.020	0.019	0.014
LOU	Mean	30.000	6.714	3.541	0.000	5.685	0.686	0.693	0.009
	±*SE*	0.000	0.459	0.203	0.000	0.682	0.020	0.018	0.019

(Continues)

TABLE 2 (Continued)

Sampled locations		*N*	N_A	N_E	P_A	A_R	H_O	H_E	F_{IS}
MAD	Mean	30.000	6.286	3.740	0.001	5.467	0.743	0.701	−0.068
	±*SE*	0.000	0.389	0.223	0.001	0.637	0.019	0.023	0.015
MUN	Mean	21.000	5.286	3.227	0.002	4.950	0.633	0.629	−0.019
	±*SE*	0.000	0.322	0.283	0.002	0.605	0.036	0.033	0.034
OLI	Mean	48.714	5.857	3.100	0.034	4.643	0.633	0.645	0.026
	±*SE*	0.143	0.442	0.203	0.034	0.576	0.026	0.021	0.018
REG	Mean	30.000	6.857	4.099	0.000	5.995	0.695	0.730	0.051
	±*SE*	0.000	0.550	0.282	0.000	0.792	0.021	0.016	0.009
REQ	Mean	29.857	5.286	3.438	0.000	4.765	0.646	0.679	0.043
	±*SE*	0.071	0.340	0.234	0.000	0.581	0.021	0.019	0.025
ROC	Mean	30.000	8.571	4.581	0.001	6.817	0.714	0.737	0.035
	±*SE*	0.000	0.635	0.415	0.001	0.953	0.026	0.021	0.010
SOP	Mean	20.000	5.857	3.318	0.000	5.400	0.621	0.649	0.041
	±*SE*	0.000	0.442	0.240	0.000	0.753	0.040	0.031	0.042
STG	Mean	50.000	9.714	5.086	0.054	6.737	0.777	0.745	−0.054
	±*SE*	0.000	0.696	0.564	0.046	1.093	0.019	0.023	0.027
VAL	Mean	49.857	4.714	3.008	0.004	4.141	0.604	0.626	0.016
	±*SE*	0.071	0.389	0.240	0.004	0.647	0.028	0.023	0.042
VIL	Mean	31.429	6.429	3.854	0.000	5.434	0.629	0.666	0.062
	±*SE*	0.101	0.606	0.415	0.000	0.891	0.035	0.032	0.012
VLB	Mean	29.000	7.286	4.223	0.001	6.162	0.739	0.726	−0.027
	±*SE*	0.000	0.574	0.324	0.001	0.830	0.020	0.022	0.020
VLR	Mean	29.714	6.143	3.404	0.000	5.034	0.654	0.659	0.005
	±*SE*	0.092	0.631	0.271	0.000	0.792	0.024	0.026	0.009
Total	Mean	32.030	6.857	3.941	0.036	5.761	0.675	0.695	0.028
	±*SE*	0.748	0.213	0.134	0.011	0.192	0.011	0.009	0.010

N = average number of crayfish in each location/loci, N_A = mean number of alleles observed, N_E = mean number of effective alleles, P_A = mean number of private alleles corrected by the sampling size, A_R = mean allelic richness, H_O = mean observed heterozygosity, H_E = mean expected heterozygosity, and F_{IS} = mean fixation index. Values are shown by mean of the seven microsatellite markers and *SE*.

Donnelly, 2000). This Bayesian clustering method assigns individuals to a given number of genetic clusters (*K*) based on their genotypes. To identify the number of clusters, we first analysed the likelihood of models with a number of clusters ranging from *K* = 1 to 27 (*n* – 1). Due to the large number of clusters, we performed 20 independent runs for each *K*, each run involving a Markov chain Monte Carlo using 2,000 burn-in followed by 10,000 iteration steps. Once preliminary results were obtained and to get more accuracy, another analysis was performed from *K* = 1 to 8 with 20 independent runs for each K, each run involving a Markov chain Monte Carlo using 200,000 burn-in followed by 1,000,000 iteration steps. Admixture ancestry models and correlated allele frequencies (with default parameters) were considered in all cases. The most likely value of real number of clusters in the genetic dataset was estimated by examining the log probability of data [Ln Pr(*X*|*K*)] and the Δ*K* method (Evanno, Regnaut, & Goudet, 2005) using STRUCTURE HARVESTER (Earl & vonHoldt, 2012). We summarised the clustering results of multiple runs for each K value and these were visually evaluated in CLUMPAK (http://clumpak.tau.ac.il) (Kopelman, Mayzel, Jakobsson, Rosenberg, & Mayrose, 2015). Additionally, a discriminant analysis of principal components (DAPC) was performed to identify the number of different clusters without assuming marker linkage neither HWE (Jombart, Devillard, & Balloux, 2010). This multivariate method consists of a two-step procedure to characterise population subdivision, being a principal component analysis as a prior step to discriminant analysis (Jombart et al., 2010). The DAPC was performed using *adegenet* version 2.1.1 (Jombart, 2008) in the R environment.

2.3 | Historical, human, and environmental drivers of genetic diversity

To introduce the historical factor in our models, we firstly used the grouping of locations resulting from STRUCTURE (Figure 1; Table S4) to generate a new categorical variable (genetic group) with two levels (Badajoz and Lower Guadalquivir). Genetic group thus identifies the original introduction foci that originated each of the present-day populations.

We then evaluated two alternative scenarios of human-driven spread through secondary introductions: the jump-dispersal scenario and the invasion hub scenario. To test the jump-dispersal scenario (Figure 2a), we calculated the linear distance (in km) of each location to its corresponding introduction foci (Badajoz [BDJ] and Lower Guadalquivir [LGQ], based on STRUCTURE results) and used this variable (step distance) as a continuous predictor of genetic diversity. This scenario assumes that transport distances are relatively constant among secondary introductions, resulting in an increasing number of jumps for increasing distances. By contrast, to test the invasion hub scenario, we selected the Ebro Delta and Valencia Albufera (northeast and east coast of Iberian Peninsula, respectively; see Figure 2b) as plausible invasion hubs, because both are large coastal wetlands with vast areas devoted to rice cropping (similar to the two original introduction foci), which received an important amount of the red swamp crayfish soon after the initial introduction (1978 in the Albufera and 1979 in the Ebro Delta; Gutierrez-Yurrita et al., 1999) and where the species have reached high densities (e.g. Clavero, López, Franch, Pou-Rovira, & Queral, 2015). To test the plausibility of the invasion hub scenario, we calculated the distance (in km) of each location to their associated introduction foci (same as for step distance) and to the candidate invasion hubs (Ebro Delta and Valencia Albufera), and selected the minimum value among these distances to generate a new continuous predictor (hub distance) of genetic diversity.

Finally, we characterised the environmental suitability for the red swamp crayfish in each location, in order to test whether higher levels of genetic diversity were related to higher suitability values. We obtained the estimated suitability based on the results of the species distribution model presented by Capinha and Anastácio (2011). These authors collected red swamp crayfish records worldwide, including native and non-native areas, and used six climatic (annual mean temperature, mean temperature of warmest quarter, mean temperature of coldest quarter, annual precipitation, precipitation of wettest quarter, and precipitation of driest quarter) and four physiographic (altitude, slope, distance to ocean, and a compound topographical index) variables to predict the species occurrence in the Iberian Peninsula with a cell resolution of 1 × 1 km. The environmental suitability, which we used as a continuous predictor of genetic diversity, was calculated as the average value of the 1 × 1 km cells included within a 5-km buffer constructed around each of our locations, excluding sea surface whenever it was included inside the buffer (Figure S1).

2.4 | Statistical analyses

We used generalised linear models (GLMs) to test the influence of the historical, human and environmental factors on the spatial patterns of genetic diversity of red swamp crayfish in the Iberian Peninsula. We ran GLMs using two genetic diversity indices (A_R and H_E) as dependent variables and genetic group, step distance, hub distance, environmental suitability, and sampling size (i.e. number of individuals analysed in each location) as predictors. Generalised linear models used normal error distribution and identity link function for both dependent variables (A_R, Shapiro–Wilk, $W = 0.98$, $p = 0.89$; H_E, Shapiro–Wilk, $W = 0.97$, $p = 0.60$). We first ran univariate GLMs testing the influence of each of the five predictors on each of the two dependent variables. Then, we ran multivariate GLMs and selected final models following a backward stepwise procedure, through which predictors were sequentially excluded from the models attending at the significance of their effects (i.e. higher *p*-values excluded first) until all predictors had either significant or marginally significant *p*-values (i.e. ≤0.1). Backward stepwise procedures for variable selection in multiple regression-type models often use $p = 0.1$ as a threshold to retain or remove variables. When forward and backward procedures are combined, a common strategy is to use $p < 0.05$ to enter and $p > 0.1$ to remove (e.g. Swartz,

Hossack, Muths, Newell, & Lowe, 2019). Generalised linear models were conducted with the *lme4* package (Bates et al., 2015) in R v3.2.3 (R Development Core Team, 2014).

3 | RESULTS

3.1 | Genetic diversity

We genotyped eight polymorphic microsatellite loci for 903 red swamp crayfish specimens from 28 locations distributed across the Iberian Peninsula. The PclG-29 locus was discarded from our dataset because it had evidence of null alleles and scoring errors due to stuttering. We thus carried out subsequent analyses with the remaining seven loci. The seven microsatellite loci exhibited moderate to high levels of polymorphism (H_E between 0.56 and 0.79) across all locations (Table 2). Most microsatellite loci were found to be in HWE, except locus PCSH02 in Ança and BDJ locations, PCSH06 in BDJ, PCSH65 in Ança and BDJ, PclG-15 in the BDJ and Guadiamar and PclG-17 in the BDJ location (Table S2). Badajoz location presented five out of seven loci in Hardy–Weinberg disequilibrium and also six locus comparisons with significant linkage disequilibrium.

We found a total of 98 alleles in the seven microsatellite loci genotyped, with polymorphism ranging from 21 (PclG-15) to eight alleles (PCSH38). At population level, the average number of alleles per locus (N_A) ranged between 4.71 ± 0.78 *SE* in Valle location (VAL), and 10.57 ± 1.54 *SE* in LGQ, where a large number of crayfish was introduced. Overall, measures of genetic diversity such as A_R, H_O, and H_E were relatively high (Table 2). The A_R ranged from 4.14 (VAL) to 7.93 (LGQ), H_O ranged from 0.51 (Jaén [JAE]) to 0.79 (LGQ) and H_E varied from 0.60 (JAE) to 0.80 (LGQ). While LGQ location had the highest genetic diversity among all locations, with the highest allelic richness (7.93 ± 2.65, mean ± *SE*), observed (0.79 ± 0.04) and expected heterozygosity values (0.80 ± 0.04), the BDJ location had low levels of genetic diversity and the highest F_{IS} value (0.22 ± 0.06) across all locations.

We tested whether any of 28 locations had recently experienced a population bottleneck. According to the Wilcoxon's test with two tails, while the TPM model showed that 14 locations had probably experienced bottleneck ($p < 0.05$), the SMM model indicated that only BDJ location could have experienced bottleneck ($p < 0.05$), although other four locations were marginally significant ($0.05 < p > 0.10$; Table S3). The BDJ location was the only one that had probably experienced a bottleneck using both models.

The BDJ location showed the lowest genetic diversity values and highest F_{IS} values ($F_{IS} = 0.22$) among all locations, lower observed than expected heterozygosity, frequent significant Hardy–Weinberg deviations and likely genetic bottleneck (Table 2 and Table S3). This fact could be explained by the origin of these samples, which were collected from aquaculture ponds (where crayfish were originally introduced) that could have been largely isolated from free-ranging crayfish for an unknown period of time. Because of this, we used the BDJ location for the analyses related to the genetic structure, but not for analyses of genetic diversity patterns.

3.2 | Genetic structure

Clustering genetic structure analysis showed that the most solid structure of red swamp crayfish in the Iberian Peninsula is the one assuming three distinct genetic clusters, $K = 3$ (Figure 1 and Figure S2). Six locations were mainly assigned to cluster 1 (orange), which mostly included Portuguese locations, as well as the Spanish BDJ and JAE locations. This cluster arguably corresponds to the group of locations originated by the spread of the crayfish introduced to Badajoz in 1973, and we henceforth refer to it as the Badajoz group (Table S4). Sixteen locations were assigned to cluster 2 (blue) and the remaining six locations were mainly included in cluster 3 (purple). Clusters 2 and 3 grouped several and widespread Iberian locations including the Lower Guadalquivir area, where a large batch of red swamp crayfish was introduced in 1974. The identification of cluster 1 (i.e. the Badajoz group) remained constant for different K values (see Figure 1 for $K = 2$, $K = 3$ and $K = 6$), while clusters 2 and 3 were grouped in a single cluster for $K = 2$. As we found no clear geographical structure between cluster 2 and 3, we consider them together as the Lower Guadalquivir group. The DAPC analysis grouped the locations in concordance with the STRUCTURE results, with a first axis separating the locations belonging either to the Lower Guadalquivir or Badajoz groups and a second axis that separates mainly the two clusters of the Lower Guadalquivir group (cluster 2 and 3) (Figure S3). STRUCTURE and DAPC analyses both supporting an admixed origin of the Mundo location, which was intermediate between Badajoz and Lower Guadalquivir group. However, it is noteworthy that, unlikely STRUCTURE analysis, the Brugent location was not grouped with Lower Guadalquivir in cluster 2.

For the isolation by distance analyses, we did not find any relationship between geographic and genetic distances among red swamp crayfish locations in the Iberian Peninsula, neither when considering all analysed locations nor when analysing the Badajoz and Lower Guadalquivir group independently (Mantel tests, $p > 0.5$ in all cases).

3.3 | Drivers of genetic diversity

The GLMs showed that the hub distance (i.e. the minimum distance of one location to its corresponding introduction foci or invasion hub) had a negative influence on the two descriptors of genetic diversity used in the analyses (Table 3; Figure 3). These relationships were significant in all univariate models and were kept in all multivariate ones (Table S5). Moreover, they were evident for both Badajoz and Lower Guadalquivir genetic groups (Figure 3). No other predictor was consistently maintained in the multivariate models. For instance, the step distance (i.e. the distance of each location

TABLE 3 Univariate and multivariate general linear models assessing the influence of different predictors on the estimators of genetic diversity (allelic richness and expected heterozygosity). Results are provided in terms of the direction of the relationship (positive, POS, or negative, NEG) for continuous predictors and comparing Lower Guadalquivir (LGQ) and Badajoz (BDJ) groups, for the genetic group factor, with an indication of the statistical significance. The coefficient of determination of the final multivariate models (selected following a backward procedure) is also shown (for full models see Table S5)

		Genetic group	*N*	Step distance	Hub distance	Environmental suitability
Allelic richness	Univariate	LGQ > BDJ	POS**	NEG*	NEG**	POS
	Multivariate (R^2 = 0.54)		POS		NEG*	
Expected heterozygosity	Univariate	LGQ > BDJ	POS	NEG	NEG*	POS
	Multivariate (R^2 = 0.13)				NEG*	

*p < 0.05;
**p < 0.01.

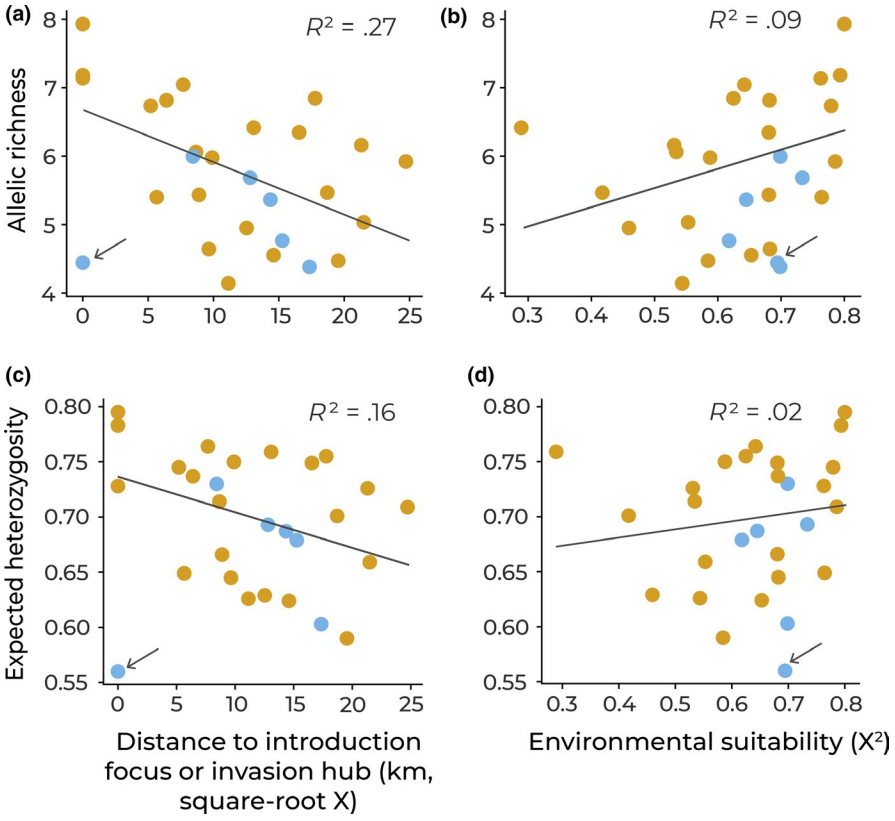


FIGURE 3 Genetic diversity indices (allelic richness and expected heterozygosity) of the 28 red swamp crayfish locations in relation to: (a, c) the minimum distance of each location to its respective introduction focus (Badajoz or Lower Guadalquivir, see Figure 1) or the closest invasion hub (Valencia Albufera or Ebro Delta); and (b, d) the environmental suitability of the area occupied by those locations. Each location is assigned to one of the two genetic groups identified (Badajoz group, blue dots; Lower Guadalquivir group, orange dots). Linear regression lines and associated coefficients of determination for the two genetic groups pooled and with the Badajoz population excluded (marked in all panels with a yellow arrow, see results) are also shown

to its corresponding introduction foci) had a weaker, often non-significant, effect on univariate models than hub distance and it was not included in any of the multivariate models (Table S5). These patterns support the existence of genetically diverse invasion hubs other than the original introduction foci, which would have served as sources for secondary introductions (i.e. invasion hub scenario, Figure 2b).

Genetic diversity figures tended to be higher in locations belonging to the Lower Guadalquivir group than in those of the Badajoz group and to be higher in areas with higher environmental suitability (Figure 3), although none of these effects were significant in univariate models (Table 3). However, the non-significance of environmental suitability could be related with the relatively high genetic diversity values of the Jiloca location, the one with the lowest suitability values (see Table 1; Figure S1).

4 | DISCUSSION

4.1 | Introduction history, genetic structure, and political boundaries

The genetic patterns observed for the red swamp crayfish in the Iberian Peninsula are associated with a complex human-mediated dispersal process involving both short- and long-distance translocations and give insight to the importance of invasion hubs, which have resulted in a lack of relationship between genetic and spatial distances. This scenario may differ from that described for other invasive freshwater organisms, when expansion after an initial introduction is due to unaided dispersal, to human-driven dispersal involving only short-distance transport or to a combination of both processes. For example, Díez-del-Molino et al. (2013) reported a

positive relationship between genetic and spatial distances for Spanish eastern mosquitofish (*Gambusia holbrooki*) populations. However, long-distance human-mediated spread of invasive species is currently a frequent feature of aquatic invasions (Audzijonyte, Baltrūnaitė, Väinölä, & Arbačiauskas, 2017; Dias et al., 2018; Marescaux et al., 2015) and has been already described for the red swamp crayfish at the global scale (Oficialdegui et al., 2019).

We identified two robust genetic groups among red swamp crayfish locations in the Iberian Peninsula, which arguably derive from the quasi-independent expansion of the two crayfish batches introduced to Spain (1973 in Badajoz and 1974 in Lower Guadalquivir). Despite the sources of crayfish for both introductions being arguably close areas in Louisiana (i.e. the native range), the lack of a strong genetic structure and the large degree of genetic admixture in Louisiana (Oficialdegui et al., 2019) could have favoured a random genetic distinction between the two transported batches due to founder effect. Once in the Iberian Peninsula, the Badajoz group would have expanded mainly westward into Portugal, but also, though less intensely, eastward (JAE location). The Lower Guadalquivir group comprised most of the red swamp crayfish Spanish range, including also a location in North-eastern Portugal (VLR location). The probability of belonging to a given group was very high around the introduction foci of both groups (i.e. near Badajoz or around the Lower Guadalquivir, Table S4), a pattern that strengthen the assumption that the observed genetic groups clusters correspond to those initial introduction events. Similarly, in a previous study based on mitochondrial DNA, Oficialdegui et al. (2019) found one haplotype (Hap_06) that was present in most Portuguese locations, but was not detected in the Lower Guadalquivir basin.

Since most alien species in inland waters are dispersed by human vectors (Cerri, Ciappelli, Lenuzza, Zaccaroni, & Nocita, 2018; Strayer, 2010), clear spatial structuring of genetic variability is often lacking among populations of invasive freshwater species (Audzijonyte et al., 2017; Blakeslee et al., 2017). However, we observed a strong genetic structure among red swamp crayfish populations in the Iberian Peninsula, probably generated by the two expansion ways from both introduction foci. A similar pattern was observed in the invasion process of the European green crab in North America, in which two separately introduction events led to two genetically distinct groups (Jeffery et al., 2017). Although natural and artificial barriers in rivers can define the spatial distribution or expansion of freshwater invasive species (see Teixeira, Neto, Gomes, Beheregaray, & Carvalho, 2020), the limited admixture in both genetic groups of red swamp crayfish (Badajoz and Lower Guadalquivir) suggests an effect related to the political border between Spain and Portugal. The border has apparently favoured the existence of two quasi-independent expansion processes, despite both countries sharing several river basins through which natural dispersion of invasive species may occur (Gago, Anastácio, Gkenas, Banha, & Ribeiro, 2016). In fact, the red swamp crayfish may have entered in Portugal through natural dispersion since the first record in the country is very near to the introduction area in Badajoz (Ramos & Pereira, 1981), although short-distance human transport cannot be discarded. However, present-day genetic patterns suggest that most subsequent human-driven translocations have remained within the political border of Portugal, with few additional introductions from Spain (even though at least one other did occur, see VLR location). Contrastingly, the expansion of the red swamp crayfish across Spain relied on the transport of individuals belonging mainly to the Guadalquivir group (except JAE location). The genetic structure within the Lower Guadalquivir group did not follow any clear spatial pattern (absence of isolation by distance and spatial distribution of clusters 2 and 3), fitting well with the patterns reported for other widely spread freshwater species (see above). The political border between Spain and Portugal may thus act as an actual ecological barrier, as has been already described for other taxa (Arrondo et al., 2018; García et al., 2018). In the red swamp crayfish case, this barrier does not seem to be related to policy differences between countries (as reported Arrondo et al., 2018), but to the behaviour of the people stocking crayfish, who apparently tended to remain within national limits. We thus emphasise the importance of political boundaries as invisible barriers that can determine the structure of wild populations, especially so for those species translocated by humans, calling for an international coordination management measures accordingly (Dresser, Pierson, & Fitzpatrick, 2018; Rollins, Woolnough, Wilton, Sinclair, & Sherwin, 2009).

4.2 | Drivers of genetic diversity

Propagule pressure is a key factor modulating the probability of establishment in introduced populations and their dynamics of genetic diversity (Lockwood, Cassey, & Blackburn, 2005). Overall, genetic diversity indices were lower in locations of the Badajoz group than in those of the Lower Guadalquivir group, a pattern probably related to the higher propagule size of the introduction event into the Lower Guadalquivir. In fact, the number of crayfish involved in the Guadalquivir introduction was around 20 times larger than in that in Badajoz (Habsburgo-Lorena, 1978).

The nature of transport events (short-distance, long-distance or both) and the existence of one or several invasion hubs acting as genetically diverse sources of individuals (e.g. Figure 2a,b) may influence the genetic diversity spatial patterns of an already established invasive species. In the Iberian Peninsula, it seems that large stocks of red swamp crayfish specimens were long-distance translocated without intermediary bottlenecks (Gutierrez-Yurita et al., 1999), generating highly genetically diverse invasion hubs that subsequently acted as source for multiple secondary introduction events. Accordingly, we found that genetic diversity tended to decrease in locations that were farther from either its respective introduction focus or invasion hub. This steady decline in genetic diversity along the invasion process corroborate the genetic consequences described for population expansion range (Austerlitz et al., 1997; Excoffier et al., 2009). For example, White, Perkins, Heckel, and Searle (2013) found a significant decline in genetic diversity across the expansion range of the bank vole (*Myodes glareolus*) in Ireland. Based on historical information and environmental characteristics, we had selected a priori the Ebro Delta

and the Valencia Albufera as plausible invasion hubs, but additional invasion hubs could have existed. Candidate areas for this role could be the Tablas de Daimiel National Park, where the red swamp crayfish was introduced in 1982, or the northern Spanish plateau, where a long-lasting tradition around crayfish consumption could have favoured the occurrence of several introduction events (Clavero, 2016). Detecting possible invasion hubs is a determining factor to predict and prevent potential range expansion of invasive species over non-native territories (Muirhead & MacIsaac, 2005).

How the genetic diversity is affected by the distance to the introduction focus and by the accumulation of bottleneck effects have been well studied in invasion biology (van Boheemen et al., 2017). However, our results also showed a positive relationship between genetic diversity and environmental suitability for the red swamp crayfish in the Iberian Peninsula, highlighting the importance of ecological factors in shaping the genetic patterns of invasive species. Similar patterns have been reported for other invertebrate species (Ortego, Aguirre, Noguerales, & Cordero, 2015). They support the influence of environmental suitability on evolutionary processes through demographic mechanisms that affect the effective population size (Wang, 2012) and ultimately modulate the genetic patterns of populations. Furthermore, the relevance of the suitability–genetic diversity relationships for the management of biological invasions can be modulated by climate change. Capinha, Anastácio, and Tenedório (2012) predicted that suitable areas for the red swamp crayfish in the Iberian Peninsula would show moderate changes (including both suitability increases and decreases) from the present situation. However, warmer future environments will enhance climate suitability for the red swamp crayfish across several European areas (Zhang et al., 2020), which according to our results, could end up hosting viable and more genetically diverse red swamp crayfish populations.

4.3 | Conclusions

We have described clear patterns in genetic structure of red swamp crayfish in the Iberian Peninsula, determined by the two introduction events that took place in the 1970s, a complex human-mediated dispersal process involving the presence of invasion hubs, and a tendency of higher levels of genetic diversity occurring in more suitable environments. These results help to comprehend the invasion history of the red swamp crayfish in the Iberian Peninsula and how the natural and anthropogenic factors modulate it. Our study thus highlights the importance of analysing patterns of genetic variability to understand the invasion processes, a knowledge that can be applied to manage current invasions and prevent possible future invasions.

ACKNOWLEDGMENTS

We thank R. López, M. A. Bravo, and Á. Vidal for field and lab assistance, as well as the many colleagues who kindly provided samples from Spain—A. Pradilla (CCEDCV - EL PALMAR), L. M. Álvarez (Asturias Government), N. Franch (Parc Natural del Delta de l'Ebre), L. Valls, F. J. Galindo (Andalusian Government), F. Unzu and L. Lopo (La Rioja Government), J. Bosch, F. J. Oliva, I. Vedia—and Portugal—F. Ribeiro and A. F. Filipe. We also thank Laboratorio de Ecología Molecular (LEM-EBD) at Estación Biológica de Doñana (CSIC) as well as C. Lejeusne and C. Daguin-Thiébaut for their advices and help in fine-tuning of the multiplex. We are especially grateful to César Capinha for sharing his dataset with us. We thank two anonymous reviewers who gave helpful comments to improve the final version of this manuscript. F.J.O. was supported by an Andalusian Government grant (P12-RNM-936). This study was funded by the Andalusian Government (RNM-936) and PICS (PIC2015FR4 and PICS07360).

ORCID

Lucía Acevedo-Limón https://orcid.org/0000-0002-0538-1362
Francisco J. Oficialdegui https://orcid.org/0000-0001-6223-736X
Marta I. Sánchez https://orcid.org/0000-0002-8349-5410
Miguel Clavero https://orcid.org/0000-0002-5186-0153

REFERENCES

Aljanabi, S. (1997). Universal and rapid salt-extraction of high quality genomic DNA for PCR-based techniques. *Nucleic Acids Research*, *25*(22), 4692–4693. https://doi.org/10.1093/nar/25.22.4692

Arrondo, E., Moleón, M., Cortés-Avizanda, A., Jiménez, J., Beja, P., Sánchez-Zapata, J. A., & Donázar, J. A. (2018). Invisible barriers: Differential sanitary regulations constrain vulture movements across country borders. *Biological Conservation*, *219*, 46–52. https://doi.org/10.1016/j.biocon.2017.12.039

Audzijonyte, A., Baltrūnaitė, L., Väinölä, R., & Arbačiauskas, K. (2017). Human-mediated lineage admixture in an expanding Ponto-Caspian crustacean species *Paramysis lacustris* created a novel genetic stock that now occupies European waters. *Biological Invasions*, *19*(8), 2443–2457. https://doi.org/10.1007/s10530-017-1454-9

Austerlitz, F., Jung-Muller, B., Godelle, B., & Gouyon, P. H. (1997). Evolution of coalescence times, genetic diversity and structure during colonization. *Theoretical Population Biology*, *51*(2), 148–164.

Banks, S. C., Cary, G. J., Smith, A. L., Davies, I. D., Driscoll, D. A., Gill, A. M., … Peakall, R. (2013). How does ecological disturbance influence genetic diversity? *Trends in Ecology and Evolution*, *28*(11), 670–679. https://doi.org/10.1016/j.tree.2013.08.005

Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H., Rcpp, L. (2015). Package 'lme4'. *Convergence*, 12(1).

Belfiore, N. M., & May, B. (2000). Variable microsatellite loci in red swamp crayfish, *Procambarus clarkii*, and their characterization in other crayfish taxa. *Molecular Ecology*, *9*, 2231–2234. https://doi.org/10.1046/j.1365-294X.2000.105339.x

Bellard, C., Cassey, P., & Blackburn, T. M. (2016). Alien species as a driver of recent extinctions. *Biology Letters*, *12*(2), 20150623. https://doi.org/10.1098/rsbl.2015.0623

Blackburn, T. M., & Duncan, R. P. (2001). Determinants of establishment success in introduced birds. *Nature*, *414*(6860), 195–197. https://doi.org/10.1038/35102557

Blackburn, T. M., Prowse, T. A. A., Lockwood, J. L., & Cassey, P. (2013). Propagule pressure as a driver of establishment success in deliberately introduced exotic species: Fact or artefact? *Biological Invasions*, *15*(7), 1459–1469. https://doi.org/10.1007/s10530-013-0451-x

Blackburn, T. M., Pyšek, P., Bacher, S., Carlton, J. T., Duncan, R. P., Jarošík, V., … Richardson, D. M. (2011). A proposed unified framework for biological invasions. *Trends in Ecology and Evolution*, *26*(7), 333–339. https://doi.org/10.1016/j.tree.2011.03.023

Blakeslee, A. M. H., Kamakura, Y., Onufrey, J., Makino, W., Urabe, J., Park, S., … Miura, O. (2017). Reconstructing the invasion history of the Asian shorecrab, *Hemigrapsus sanguineus* (De Haan 1835) in the

Western Atlantic. *Marine Biology*, *164*, 47. https://doi.org/10.1007/s00227-017-3069-1

Bossdorf, O., Richards, C. L., & Pigliucci, M. (2008). Epigenetics for ecologists. *Ecology Letters*, *11*(2), 106–115. https://doi.org/10.1111/j.1461-0248.2007.01130.x

Capinha, C., & Anastácio, P. (2011). Assessing the environmental requirements of invaders using ensembles of distribution models. *Diversity and Distributions*, *17*(1), 13–24. https://doi.org/10.1111/j.1472-4642.2010.00727.x

Capinha, C., Anastácio, P., & Tenedório, J. A. (2012). Predicting the impact of climate change on the invasive decapods of the Iberian inland waters: An assessment of reliability. *Biological Invasions*, *14*(8), 1737–1751. https://doi.org/10.1007/s10530-012-0187-z

Capinha, C., Essl, F., Seebens, H., Moser, D., & Pereira, H. M. (2015). The dispersal of alien species redefines biogeography in the Anthropocene. *Science*, *348*(6240), 1248–1251. https://doi.org/10.1126/science.aaa8913

Cerri, J., Ciappelli, A., Lenuzza, A., Zaccaroni, M., & Nocita, A. (2018). Recreational angling as a vector of freshwater invasions in Central Italy: Perceptions and prevalence of illegal fish restocking. *Knowledge & Management of Aquatic Ecosystems*, *419*, 38. https://doi.org/10.1051/kmae/2018028

Clavero, M. (2016). Species substitutions driven by anthropogenic positive feedbacks: Spanish crayfish species as a case study. *Biological Conservation*, *193*, 80–85. https://doi.org/10.1016/j.biocon.2015.11.017

Clavero, M., López, V., Franch, N., Pou-Rovira, Q., & Queral, J. M. (2015). Use a seasonally flooded rice fields by fish and crayfish in a Mediterranean wetland. *Agriculture, Ecosystems & Environment*, *213*, 39–46. https://doi.org/10.1016/j.agee.2015.07.022

Dias, P. J., Gilg, M. R., Lukehurst, S. S., Kennington, W. J., Huhn, M., Madduppa, H. H., … McDonald, J. I. (2018). Genetic diversity of a hitchhiker and prized food source in the Anthropocene: The Asian green mussel *Perna viridis* (Mollusca, Mytilidae). *Biological Invasions*, *20*(7), 1749–1770.

Díez-del-Molino, D., Carmona-Catot, G., Araguas, R. M., Vidal, O., Sanz, N., García-Berthou, E., & García-Marín, J. L. (2013). Gene flow and maintenance of genetic diversity in invasive mosquitofish (*Gambusia holbrooki*). *PLoS ONE*, *8*(12), e82501. https://doi.org/10.1371/journal.pone.0082501

Dlugosch, K. M., & Parker, I. M. (2008). Founding events in species invasions: Genetic variation, adaptive evolution, and the role of multiple introductions. *Molecular Ecology*, *17*(1), 431–449. https://doi.org/10.1111/j.1365-294X.2007.03538.x

Dray, S., & Dufour, A.-B. (2007). The ade4 package: Implementing the duality diagram for ecologists. *Journal of Statistical Software*, *22*(4), 1–20. https://doi.org/10.18637/jss.v022.i04.

Dresser, C. M., Pierson, T. W., & Fitzpatrick, B. M. (2018). Isolation by distance, local adaptation, and fortuitous coincidence of geo-political boundaries with spatial-genetic clusters in southern Bog Turtles. *Global Ecology and Conservation*, *16*, e00474. https://doi.org/10.1016/j.gecco.2018.e00474

Drolet, D., & Locke, A. (2016). Relative importance of propagule size and propagule number for establishment of non-indigenous species: A stochastic simulation study. *Aquatic Invasions*, *11*(1), 101–110. https://doi.org/10.3391/ai.2016.11.1.11

Earl, D. A., & vonHoldt, B. M. (2012). STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, *4*(2), 359–361. https://doi.org/10.1007/s12686-011-9548-7

Ellegren, H., & Galtier, N. (2016). Determinants of genetic diversity. *Nature Reviews Genetics*, *17*(7), 422–433. https://doi.org/10.1038/nrg.2016.58

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Molecular Ecology*, *14*(8), 2611–2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x

Excoffier, L., Foll, M., & Petit, J. (2009). Genetic consequences of range expansions. *Annual Review of Ecology, Evolution, and Systematics*, *40*, 481–501. https://doi.org/10.1146/annurev.ecolsys.39.110707.173414

Excoffier, L., Laval, G., & Schneider, S. (2005). Arlequin (version 3.0): An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics*, *1*, 47–50. https://doi.org/10.1177/117693430500100003

Facon, B., Genton, B. J., Shykoff, J., Jarne, P., Estoup, A., & David, P. (2006). A general eco-evolutionary framework for understanding bioinvasions. *Trends in Ecology and Evolution*, *21*(3), 130–135. https://doi.org/10.1016/j.tree.2005.10.012

Gago, J., Anastácio, P., Gkenas, C., Banha, F., & Ribeiro, F. (2016). Spatial distribution patterns of the non-native European catfish, *Silurus glanis*, from multiple online sources – A case study for the River Tagus (Iberian Peninsula). *Fisheries Management and Ecology*, *23*(6), 503–509. https://doi.org/10.1111/fme.12189

Geiger, W., Alcorlo, P., Baltanas, A., & Montes, C. (2005). Impact of an introduced Crustacean on the trophic webs of Mediterranean wetlands. *Biological Invasions*, *7*(1), 49–73. https://doi.org/10.1007/1-4020-3870-4_6

Gherardi, F. (2006). Crayfish invading Europe: The case study of *Procambarus clarkii*. *Marine and Freshwater Behaviour and Physiology*, *39*(3), 175–191. https://doi.org/10.1080/10236240600869702

Gutiérrez-Yurrita, P. J., Martinez, J. M., Ilhéu, M., Bravo-Utrera, M. A., Bernardo, J. M., & Montes, C. (1999). The status of crayfish populations in Spain and Portugal. *Crustacean Issues*, *11*, 161–192.

Habsburgo-Lorena, A. S. (1978). Present situation of exotic species of crayfish introduced into Spanish continental water. *Freshwater Crayfish*, *4*, 175–184.

Hardesty, B. D., Le Roux, J. J., Rocha, O. J., Meyer, J. Y., Westcott, D., & Wieczorek, A. M. (2012). Getting here from there: Testing the genetic paradigm underpinning introduction histories and invasion success. *Diversity and Distributions*, *18*(2), 147–157. https://doi.org/10.1111/j.1472-4642.2011.00832.x

Hawes, N. A., Fidler, A. E., Tremblay, L. A., Pochon, X., Dunphy, B. J., & Smith, K. F. (2018). Understanding the role of DNA methylation in successful biological invasions: A review. *Biological Invasions*, *20*(9), 2285–2300.

Hijmans, R. J., Williams, E., & Vennes, C. (2017). Package 'geosphere' version 1.5-7. Retrieved from https://CRAN.R-project.org/package=geosphere.

Hulme, P. E., Bacher, S., Kenis, M., Klotz, S., Kühn, I., Minchin, D., … Vilà, M. (2008). Grasping at the routes of biological invasions: A framework for integrating pathways into policy. *Journal of Applied Ecology*, *45*(2), 403–414. https://doi.org/10.1111/j.1365-2664.2007.01442.x

Jeffery, N. W., Dibacco, C., Wyngaarden, M. V., Hamilton, L. C., Stanley, R. R. E., Bernier, R., … Bradbury, I. R. (2017). RAD sequencing reveals genomewide divergence between independent invasions of the European green crab (*Carcinus maenas*) in the Northwest Atlantic. *Ecology and Evolution*, *7*(8), 2513–2524. https://doi.org/10.1002/ece3.2872

Jiang, H., Qian, Z., Lu, W., Xing, Z., Yu, H., & Li, J. (2015). Microsatellite marker identification from transcriptome derived sequences of the red swamp crawfish, *Procambarus clarkii*. *Conservation Genetics Resources*, *7*(3), 729–731. https://doi.org/10.1007/s12686-015-0447-1

Jombart, T. (2008). adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics*, *24*(11), 1403–1405. https://doi.org/10.1093/bioinformatics/btn129

Jombart, T., Devillard, S., & Balloux, F. (2010). Discriminant analysis of principal components: A new method for the analysis of

genetically structured populations. *BMC Genetics*, *11*(1), 94. https://doi.org/10.1186/1471-2156-11-94

Kopelman, N. M., Mayzel, J., Jakobsson, M., Rosenberg, N. A., & Mayrose, I. (2015). Clumpak: A program for identifying clustering modes and packaging population structure inferences across K. *Molecular Ecology Resources*, *15*(5), 1179–1191. https://doi.org/10.1111/1755-0998.12387

Lavergne, S., & Molofsky, J. (2007). Increased genetic variation and evolutionary potential drive the success of an invasive grass. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(10), 3883–3888. https://doi.org/10.1073/pnas.0607324104

Lockwood, J. L., Cassey, P., & Blackburn, T. (2005). The role of propagule pressure in explaining species invasions. *Trends in Ecology and Evolution*, *20*(5), 223–228. https://doi.org/10.1016/j.tree.2005.02.004

Marescaux, J., von Oheimb, K. C., Etoundi, E., von Oheimb, P. V., Albrecht, C., Wilke, T., & Van Doninck, K. (2015). Unravelling the invasion pathways of the quagga mussel (*Dreissena rostriformis*) into Western Europe. *Biological Invasions*, *18*(1), 245–264. https://doi.org/10.1007/s10530-015-1005-1

Muirhead, J. R., & MacIsaac, H. J. (2005). Development of inland lakes as hubs in an invasion network. *Journal of Applied Ecology*, *42*, 80–90. https://doi.org/10.1111/j.1365-2664.2004.00988.x

Oficialdegui, F. J., Green, A. J., Clavero, M., Sánchez, M. I., Boyero, L., Michot, T. C., … Lejeusne, C. (2019). Unravelling the global invasion routes of a worldwide invader, the red swamp crayfish (*Procambarus clarkii*). *Freshwater Biology*, *64*(8), 1382–1400. https://doi.org/10.1111/fwb.13312

Oficialdegui, F. J., Sánchez, M. I., & Clavero, M. (2020) One century away from home: How the red swamp crayfish took over the world. *Reviews in Fish Biology and Fisheries*, *30*(1), 121–135. https://doi.org/10.1007/s11160-020-09594-z

Okada, M., Lyle, M., & Jasieniuk, M. (2009). Inferring the introduction history of the invasive apomictic grass *Cortaderia jubata* using microsatellite markers. *Diversity and Distributions*, *15*(1), 148–157. https://doi.org/10.1111/j.1472-4642.2008.00530.x

Ortego, J., Aguirre, M. P., Noguerales, V., & Cordero, P. J. (2015). Consequences of extensive habitat fragmentation in landscape-level patterns of genetic diversity and structure in the Mediterranean esparto grasshopper. *Evolutionary Applications*, *8*(6), 621–632. https://doi.org/10.1111/eva.12273

Peakall, R., & Smouse, P. E. (2012). GenAlEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research-an update. *Bioinformatics*, *28*(19), 2537–2539 https://doi:10.1093/bioinformatics/bts460.

Piry, S., Luikart, G., & Cornuet, J. M. (1999). BOTTLENECK: A computer program for detecting recent reductions in the effective population size using allele frequency data. *Journal of Heredity*, *90*, 502–503. https://doi.org/10.1093/jhered/90.4.502

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, *155*(2), 945–959. https://doi.org/10.1111/j.1471-8286.2007.01758.x

R Development Core Team. (2014). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Ramos, M. A., & Pereira, T. G. (1981). Um novo Astacidae para a Fauna Portuguesa *Procambarus clarkii* (Girard, 1852). Boletim do Instituto Nacional de Investigação das Pescas. *Lisboa*, *6*, 37–47.

Ricciardi, A. (2007). Are modern biological invasions an unprecedented form of global change? *Conservation Biology*, *21*(2), 329–336. https://doi.org/10.1111/j.1523-1739.2006.00615.x

Rice, W. R. (1989). Analyzing tables of statistical test. *Evolution*, *43*(1), 223–225. https://doi.org/10.1111/j.0014-3820.2001.tb00731.x

Rius, M., & Darling, J. A. (2014). How important is intraspecific genetic admixture to the success of colonising populations? *Trends in Ecology and Evolution*, *29*(4), 233–242. https://doi.org/10.1016/j.tree.2014.02.003

Rollins, L. A., Woolnough, A. P., Wilton, A. N., Sinclair, R., & Sherwin, W. B. (2009). Invasive species can't cover their tracks: Using microsatellites to assist management of starling (*Sturnus vulgaris*) populations in Western Australia. *Molecular Ecology*, *18*(8), 1560–1573. https://doi.org/10.1111/j.1365-294X.2009.04132.x

Rousset, F. (2008). GENEPOP'007: A complete re-implementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources*, *8*(1), 103–106. https://doi.org/10.1111/j.1471-8286.2007.01931.x

Salanti, G., Amountza, G., Ntzani, E. E., & Ioannidis, J. P. A. (2005). Hardy-Weinberg equilibrium in genetic association studies: An empirical evaluation of reporting, deviations, and power. *European Journal of Human Genetics*, *13*(7), 840–848. https://doi.org/10.1038/sj.ejhg.5201410

Seebens, H., Blackburn, T. M., Dyer, E. E., Genovesi, P., Hulme, P. E., Jeschke, J. M., … Essl, F. (2017). No saturation in the accumulation of alien species worldwide. *Nature Communications*, *8*, 1–9. https://doi.org/10.1038/ncomms14435

Silva, P., López-Bao, J. V., Llaneza, L., Álvares, F., Lopes, S., Blanco, J. C., … Godinho, R. (2018). Cryptic population structure reveals low dispersal in Iberian wolves. *Scientific Reports*, *8*(1), 1–14. https://doi.org/10.1038/s41598-018-32369-3

Souty-Grosset, C., Anastácio, P. M., Aquiloni, L., Banha, F., Choquer, J., Chucholl, C., & Tricarico, E. (2016). The red swamp crayfish *Procambarus clarkii* in Europe: Impacts on aquatic ecosystems and human well-being. *Limnologica*, *58*, 78–93. https://doi.org/10.1016/j.limno.2016.03.003

Strayer, D. L. (2010). Alien species in fresh waters: Ecological effects, interactions with other stressors, and prospects for the future. *Freshwater Biology*, *55*, 152–174. https://doi.org/10.1111/j.1365-2427.2009.02380.x

Swartz, L. K., Hossack, B. R., Muths, E., Newell, R. L., & Lowe, W. H. (2019). Aquatic macroinvertebrate community responses to wetland mitigation in the Greater Yellowstone Ecosystem. *Freshwater Biology*, *64*(5), 942–953.

Szpiech, Z. A., Jakobsson, M., & Rosenberg, N. A. (2008). ADZE: A rarefaction approach for counting alleles private to combinations of populations. *Bioinformatics*, *24*(21), 2498–2504. https://doi.org/10.1093/bioinformatics/btn478

Teixeira, D. F., Neto, F. R. A., Gomes, L. C., Beheregaray, L. B., & Carvalho, D. C. (2020). Invasion dynamics of the white piranha (*Serrasalmus brandtii*) in a Neotropical river basin. *Biological Invasions*, *22*(3), 983–995. https://doi.org/10.1007/s10530-019-02138-y

van Boheemen, L. A., Lombaert, E., Nurkowski, K. A., Gauffre, B., Rieseberg, L. H., & Hodgins, K. A. (2017). Multiple introductions, admixture and bridgehead invasion characterize the introduction history of *Ambrosia artemisiifolia* in Europe and Australia. *Molecular Ecology*, *26*(20), 5421–5434. https://doi.org/10.1111/mec.14293

Van Oosterhout, C., Hutchinson, W. F., Wills, D. P. M., & Shipley, P. (2004). MICRO-CHECKER: Software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*, *4*(3), 535–538. https://doi.org/10.1111/j.1471-8286.2004.00684.x

Wallace, J. R. (2015). Interactive Mapping: Package 'IMAP'. version 1.32. Retrieved from https://CRAN.R-project.org/package=Imap.

Walsh, J. R., Carpenter, S. R., & Vander Zanden, M. J. (2016). Invasive species triggers a massive loss of ecosystem services through a trophic cascade. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(15), 4081–4085. https://doi.org/10.1073/pnas.1600366113

Wang, I. J. (2012). Environmental and topographic variables shape genetic structure and effective population sizes in the endangered Yosemite toad. *Diversity and Distributions*, *18*(10), 1033–1041. https://doi.org/10.1111/j.1472-4642.2012.00897.x

White, T. A., Perkins, S. E., Heckel, G., & Searle, J. B. (2013). Adaptive evolution during an ongoing range expansion: The invasive bank vole (*Myodes glareolus*) in Ireland. *Molecular Ecology*, *22*(11), 2971–2985. https://doi.org/10.1111/mec.12343

Wright, S. (1943). Isolation by distance. *Genetics*, *28*(2), 114.

Zhang, Z., Capinha, C., Usio, N., Weterings, R., Liu, X., Li, Y., … Yokota, M. (2020). Impacts of climate change on the global potential distribution of two notorious invasive crayfishes. *Freshwater Biology*, *65*(3), 353–365. https://doi.org/10.1111/fwb.13429

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Acevedo-Limón L, Oficialdegui FJ, Sánchez MI, Clavero M. Historical, human, and environmental drivers of genetic diversity in the red swamp crayfish (*Procambarus clarkii*) invading the Iberian Peninsula. *Freshwater Biology*. 2020;00:1–15. https://doi.org/10.1111/fwb.13513