



OPEN

# Genetic diversity and population structure of the rockpool shrimp *Palaemon elegans* based on microsatellites: evidence for a cryptic species and differentiation across the Atlantic–Mediterranean transition

Inés González-Castellano<sup>✉</sup>, Jorge González-López, Ana M. González-Tizón & Andrés Martínez-Lage<sup>✉</sup>

The rockpool shrimp *Palaemon elegans* is an ecologically important crustacean species within the European coastline fauna. In the present study, genetic diversity and population structure and connectivity were assessed by examining 21 polymorphic microsatellite loci at 13 sampling sites located along the natural distribution range. All localities showed similar levels of genetic variability. Significant deficits of heterozygosity were recorded, most likely due to the presence of null alleles. Genetic structure analyses revealed two clearly genetically distinct groups within *P. elegans* but without following any geographical or oceanographic basis. Thus, our results provided nuclear evidence for the existence of a Mediterranean cryptic species within *P. elegans*, highlighting the need to revise its taxonomic status. Regarding *P. elegans sensu stricto*, population structuring was reported across the Atlantic–Mediterranean transition area, where the Almería–Orán Front restricts the gene flow between the Atlantic and the Mediterranean population. Moreover, while population connectivity was suggested between all Mediterranean localities, some substructure was found within the Atlantic group. Canary Islands exhibited a weak but significant genetic differentiation from all Atlantic mainland localities, consistent with the isolation-by-distance pattern detected throughout the Atlantic population. Overall, all these findings provided new insights into the population biology of *P. elegans* complex.

The rockpool shrimp *Palaemon elegans* Rathke, 1837 is a crustacean decapod common in tidal rockpools, *Zostera*, *Posidonia* and *Cymodocea* sea grasses meadows and it also can be found in hypersaline lagoons and in slightly brackish water close to river mouths<sup>1</sup>. This species is characterized by its capability to adapt to highly variable environments, as it can cope with a wide range of salinities, temperatures and oxygen conditions<sup>2,3</sup>. Regarding to the native geographical distribution of *P. elegans*, it ranges from the eastern Atlantic Ocean (from Scotland and Norway to Mauritania, including the Azores, Madeira and Canary Islands) to the entire Mediterranean Sea and the Black Sea<sup>4</sup>. Unintentional introductions resulted in this shrimp inhabiting the Caspian Sea and the Aral

Departamento de Biología and Centro de Investigaciones Científicas Avanzadas (CICA), Universidade da Coruña, 15071 A Coruña, Spain. ✉email: ines.gonzalez.castellano@udc.es; andres@udc.es

Sea nowadays. Similarly, *P. elegans* was introduced in the Baltic Sea, replacing the native counterpart *P. adspersus* in that region<sup>5,6</sup>. *Palaemon elegans* has been also recently reported to occur on the northeast coast of the United States<sup>7</sup>, so is currently distributed beyond European waters. Thus, even though *P. elegans* is considered an important species within the European coastline fauna due to its broad ecological niche and ongoing geographic expansion<sup>8</sup>, population genetics analyses are scarce.

Genetic breaks occur when there is a marked genetic differentiation in a continuously distributed species<sup>9</sup>. For many years, high levels of gene flow associated with little genetic structuring were assumed for marine species with planktonic larval phases<sup>10</sup> as *P. elegans*. However, it has been proved that historical and oceanographic processes could drive genetic differentiation between marine populations, even in species with high dispersal abilities<sup>11</sup>. This is the case of the shrimp *P. elegans* since Reuschel et al.<sup>8</sup> carried out a phylogeographic analysis using two mitochondrial DNA (mtDNA) markers, revealing the existence of three haplogroups, one of them in the Atlantic localities (type I) and two from the Mediterranean localities (types II and III). In that study, a genetic break was found in the Atlantic–Mediterranean transition area, restricting the gene flow between these basins and leading to genetic differentiation between the Atlantic (type I) and the Mediterranean (type II) population. Similar phylogeographic breaks across the Atlantic–Mediterranean transition have been reported in a wide variety of species (see review in Patarnello et al.<sup>11</sup>). These genetic discontinuities are frequently attributed to the effect of different oceanographic discontinuities acting in this area, i.e. the Almería–Orán Front, the Strait of Gibraltar, the Balearic Front or the Ibiza Channel<sup>10</sup>.

Back to *P. elegans*, genetic structure was not detected by Reuschel et al.<sup>8</sup> neither within the Atlantic population (type I) nor within the Mediterranean one (type II). As for *P. elegans* type III, it was only recorded from Mediterranean localities and showed high nucleotide divergence from the other two types<sup>8</sup>. This genetic pattern was confirmed in later phylogeographic studies in the Mediterranean Sea also using mtDNA markers<sup>12,13</sup>. Reuschel et al.<sup>8</sup> proposed that type III individuals could belong to a putative cryptic species existing within *P. elegans*. Cryptic species occur when two or more species are classified as a single nominal species because they are superficially morphologically indistinguishable<sup>14</sup>. The Mediterranean Sea is characterized by a high level of biodiversity and species richness with a proportion of endemism circa 25%<sup>15</sup>, largely related to its tortuous geological history. Consequently, cryptic speciation is not considered an uncommon phenomenon in the Mediterranean Sea and molecular markers have indicated the presence of cryptic species in several taxa including shrimps<sup>16,17</sup>. Given these findings in *P. elegans*, the need for further genetic studies using polymorphic nuclear markers, such as microsatellites, to clarify its phylogeography was highlighted.

Microsatellite markers have been used in a wide range of applications in population genetics, ecological, conservation and evolutionary studies<sup>18</sup> because of its codominant nature, biparental mode of inheritance and high levels of polymorphism<sup>19</sup>. Specifically in population genetics, microsatellites are extremely valuable tools as they might reveal the existence of genetically distinct populations even in fine-scale studies<sup>20</sup>. In fact, microsatellite loci have been crucial in understanding the population structure of several shrimp species as *Penaeus vannamei*<sup>21</sup>, *Penaeus monodon*<sup>22</sup>, *Pandalus borealis*<sup>23</sup>, *Aristeus antennatus*<sup>24</sup> or *Palaemon serratus*<sup>25</sup>. Thanks to the advent of high-throughput sequencing technologies, the development of microsatellite markers in non-model species has become a rapid and cost-effective process<sup>26</sup>. Precisely, we isolated and characterized 21 novel polymorphic microsatellite loci for *Palaemon elegans* using Illumina MiSeq sequencing<sup>27</sup>.

The aim of the present study was to address for the first time the genetic diversity and population structure and connectivity of the shrimp *P. elegans* along the European coastlines, with particular attention to the Atlantic–Mediterranean transition, using microsatellite loci previously developed<sup>27</sup> to shed light about the population biology of this species.

## Results

**Genetic diversity.** A total of 400 individuals were genotyped for 21 microsatellite markers isolated and characterized in a previous work<sup>27</sup>. Thirteen coastal localities distributed along the natural distribution range of *P. elegans* were sampled. The number of individuals collected from each sampling site varied from 10 (Ebro Delta) to 50 (Collioure). No evidence of linkage disequilibrium between pairs of loci was detected across localities after the sequential Bonferroni correction. Hence, all microsatellites were considered as independent markers. Likewise, the Ewens–Watterson's neutrality test confirmed that all loci were selectively neutral. Estimators of genetic variability per locality, per locus and per locality and locus are available in Table 1, Supplementary Tables S1 and S2, respectively. Across the microsatellite loci, the number of alleles ranged from 4 (loci Pe04, Pe13 and Pe20) to 25 (locus Pe06), with a grand mean of 5.3 alleles over localities and loci. Thirty-six out of the 217 identified alleles (16.59%) were private alleles. Every microsatellite locus showed private alleles from at least one sampling site (except loci Pe02, Pe03, Pe08 and Pe13 with no private alleles), and every locality showed private alleles in at least one locus (except Lanzarote) (Table 1 and Supplementary Table S1). Average allelic richness (*AR*) across localities ranged from 2.032 (Santoña) to 2.453 (Livorno). The 21 microsatellites were highly polymorphic as the mean percentage of polymorphism was 98.53% (Table 1). The observed heterozygosity (*Ho*) over localities for each locus ranged from 0.134 (Pe09) to 0.842 (Pe14), and the expected heterozygosity (*He*) from 0.201 (Pe09) to 0.876 (Pe14) (Supplementary Table S1). Regarding the expected and observed heterozygosity for each sampling site, *Ho* ranged from 0.372 (Santoña) to 0.499 (Ebro Delta) and *He* from 0.468 (Santoña) to 0.607 (Livorno). *Ho* and *He* averaged 0.450 and 0.531, respectively (Table 1).

After the sequential Bonferroni correction, 242 out of the 273 locality-locus combinations (88.64%) showed no significant deviation from Hardy–Weinberg equilibrium (HWE). Nine out of the 21 loci (Pe03, Pe04, Pe05, Pe07, Pe12, Pe13, Pe15, Pe20 and Pe21) conformed to HWE in all localities, meanwhile the remaining 12 loci were out of HWE in at least one locality (Supplementary Table S2). Regarding localities, Collioure was the locality where more loci deviated from HWE, while in Lanzarote, Ebro Delta and Marseille no loci were significantly out

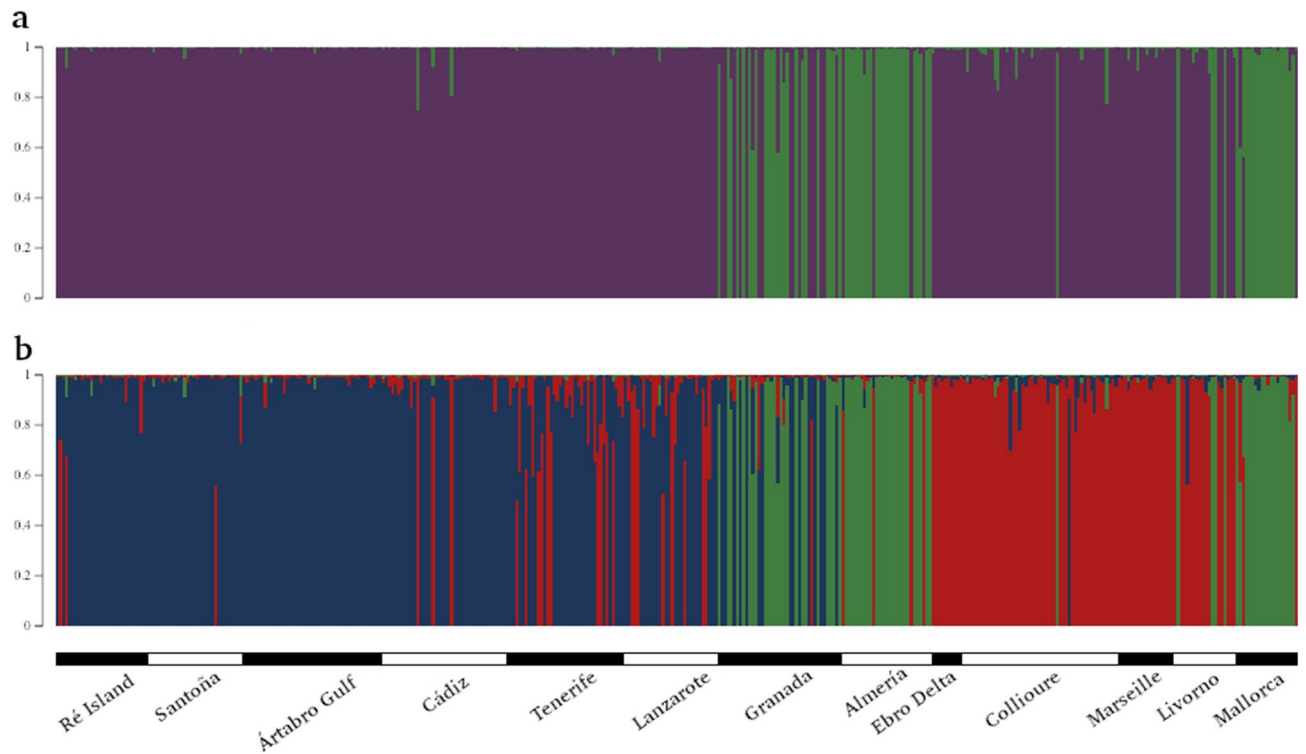
	Locality	N	Na	A	AR	P (%)	Ho	He	FIS
Atlantic Ocean	Ré Island	30	96 (1)	4.571	2.081	95.24	0.448	0.492	0.112
	Santoña	30	95 (2)	4.524	2.032	100.00	0.372	0.468	0.221
	Ártabro Gulf	45	110 (4)	5.238	2.071	100.00	0.448	0.491	0.100
	Cádiz	40	112 (5)	5.333	2.101	100.00	0.439	0.504	0.141
	Tenerife	38	111 (1)	5.286	2.178	100.00	0.481	0.529	0.104
	Lanzarote	30	103 (0)	4.905	2.186	95.24	0.483	0.528	0.103
Mediterranean Sea	Granada	40	129 (2)	6.143	2.348	100.00	0.474	0.591	0.207
	Almería	29	114 (2)	5.429	2.255	100.00	0.434	0.548	0.199
	Ebro Delta	10	95 (1)	4.524	2.298	100.00	0.499	0.539	0.127
	Collioure	50	137 (5)	6.524	2.317	100.00	0.494	0.571	0.143
	Marseille	18	106 (3)	5.048	2.219	100.00	0.422	0.518	0.212
	Livorno	20	132 (4)	6.286	2.453	100.00	0.446	0.607	0.284
	Mallorca	20	110 (6)	5.238	2.242	90.48	0.418	0.521	0.226
	Total/average	400	217 (36)	5.311	2.214	98.53	0.450	0.531	0.168

**Table 1.** Genetic diversity for 13 sampling sites of *Palaemon elegans*. N, number of sampled individuals; Na, number of alleles per locality, with private alleles in brackets; A, mean number of alleles per locality; AR, allelic richness; P, percentage of polymorphism; Ho, mean observed heterozygosity; He, mean expected heterozygosity; FIS, inbreeding coefficient.

of HWE. Most of these combinations departing from HWE were accompanied by positive FIS values, indicating the existence of a heterozygote deficit. An excess of heterozygotes from Collioure for locus Pe02 was detected and from Livorno for locus Pe10. Null allele frequencies were mostly low, varying from 0.022 to 0.226 across loci. Only one locus, the locus Pe11, exhibited a high null allele frequency, >0.2 according to Chapuis & Estoup<sup>28</sup>, over localities (Supplementary Table S1). In detail, eight loci (Pe01, Pe02, Pe06, Pe08, Pe11, Pe17, Pe18 and Pe19) showed high null allele frequency (above 0.2) in at least one locality (Supplementary Table S2), being again the locus Pe11 the one with null allele frequency >0.2 in more localities (Ré Island, Santoña, Ártabro Gulf, Tenerife, Collioure, Marseille and Livorno). Given the consistency in the presence of null alleles across localities, locus Pe11 was not included in the further population structure analyses. Specifically, 14 out of the 31 locality-locus combinations departing from HWE (48.28%) showed evidence of presence of null alleles in a frequency higher than 0.2, whilst in the remaining combinations the proportion of null alleles was lower (Supplementary Table S2).

**Genetic structure.** The assessment of the modal value of DeltaK by the STRUCTURE analysis (Supplementary Figure S1a) indicated that the most likely partition of the microsatellite dataset (400 sampled individuals) was two genetically distinct groups ( $K=2$ , Fig. 1a). Most of the individuals (327 individuals) were assigned to one cluster, whereas the second cluster grouped 69 individuals. Accordingly, only 4 out of the 400 sampled individuals could not be assigned to one specific cluster (probability of membership to a cluster <0.75), most likely because they presented missing data at several loci. This partition in two groups did not follow any geographical or oceanographic pattern because the second cluster was formed by only some of the individuals collected in Granada, Almería, Collioure, Livorno and Mallorca, being the remaining individuals from these localities located in the main cluster. Factorial correspondence analysis (FCA) also showed genetic structuring in two main groups (Supplementary Figure S2) that coincided and supported the clustering inferred by STRUCTURE. Taking into account that the smaller cluster was exclusively formed by Mediterranean individuals, coupled with the findings reported by Reuschel et al.<sup>8</sup>, we considered that this group is the putative cryptic species suggested to exist within *P. elegans* with mtDNA markers. Consequently, to check if there was such agreement between our nuclear markers and the previously analysed mtDNA markers with respect to the 69 individuals of the smaller cluster, we generated the cytochrome c oxidase subunit I (COI) sequence from them following the PCR conditions described in Reuschel et al.<sup>8</sup>. A COI fragment was successfully amplified for 64 out of the 69 individuals. BLASTn verified that the COI sequence of all these individuals corresponded to the COI sequence of those individuals named *P. elegans* type III (HE573177) and considered as a cryptic species in Reuschel et al.<sup>8</sup>. Therefore, the assignment of 69 individuals of our dataset to a putative cryptic species by the 20 microsatellite loci was also supported with a mtDNA marker. COI sequences from the individuals belonging to the putative cryptic species can be easily accessed in Supplementary File S1. In detail, distribution of the putative cryptic species across localities identified by the Bayesian analysis yielded: 22 out of the 40 individuals collected in Granada (55% sampled individuals), 25 out of the 29 individuals collected in Almería (86%), 1 out of the 50 individuals collected in Collioure (2%), 4 out of the 20 individuals collected in Livorno (20%), and 17 out of the 20 individuals collected in Mallorca (85%).

In accordance with the STRUCTURE analysis, a peak of  $\Delta K$  and a plateau in the plot of  $K$  versus  $\ln \Pr(X|K)$  (Supplementary Figure S1b) were also observed at  $K=3$ , supporting a second partition of the data in three groups (Fig. 1b). One of these groups matched with the cluster corresponding to the putative cryptic species defined in  $K=2$ . However, in this analysis, the individuals considered as *P. elegans sensu stricto* were divided into two clusters: individuals from Atlantic localities (Ré Island, Santoña, Ártabro Gulf, Cádiz, Tenerife and Lanzarote)



**Figure 1.** Bayesian assignment probabilities for (a)  $K=2$  and (b)  $K=3$  revealed by STRUCTURE based on 20 microsatellite loci. Analysis including the 400 sampled individuals. Each color depicts a cluster and each vertical bar represents an individual with the probability of membership to a cluster. Colors correspond to (a) *P. elegans sensu stricto*: purple, and cryptic species: green; and (b) *P. elegans sensu stricto* (blue and red), and cryptic species (green)."

Grouping	Source of variation	Percentage of variation	F-statistics
(a) None	Among localities	16.40	
	Within localities	83.60	$F_{ST}=0.164$
(b) <i>P. elegans s. str.</i> vs. cryptic species	Among groups	26.23	$F_{CT}=0.262$
	Among localities within groups	5.41	$F_{SC}=0.073$
	Within localities	68.36	$F_{ST}=0.316$
(c) Atlantic <i>P. elegans s. str.</i> vs. Mediterranean <i>P. elegans s. str.</i> vs. cryptic species	Among groups	20.06	$F_{CT}=0.201$
	Among localities within groups	2.69	$F_{SC}=0.034$
	Within localities	77.25	$F_{ST}=0.227$

**Table 2.** Analysis of molecular variance (AMOVA) based on different grouping hypotheses. Analyses including the 400 sampled individuals. (a) No grouping indicates that localities were analysed separately. (b) For two-group analysis, localities with *P. elegans sensu stricto* individuals (*P. elegans*) and localities with putative cryptic species individuals (cryptic species) each formed a group. (c) For three-group analysis, localities with *P. elegans sensu stricto* individuals were pooled as Ré Island-Santoña-Ártabro Gulf-Cádiz-Tenerife-Lanzarote-Granada (Atlantic *P. elegans s. str.*) and Almería-Ebro Delta-Collioure-Marseille-Livorno-Mallorca (Mediterranean *P. elegans s. str.*), meanwhile localities with putative cryptic species individuals formed a third group.

and Granada were assigned to one of them, and individuals from the Mediterranean Sea (Almería, Ebro Delta, Collioure, Marseille, Livorno and Mallorca) to the other.

When all sampling sites were analysed separately, a hierarchical locus-by-locus AMOVA (Table 2a) revealed that the proportion of global genetic variance was mostly attributed to differences among individuals within localities (83.60%), whereas the proportion of genetic variance among sampling sites was 16.40% ( $F_{ST}=0.164$ ). More AMOVA analyses were performed to test the grouping scenarios inferred by STRUCTURE. Granada, Almería, Collioure, Livorno and Mallorca localities, with individuals belonging to the putative cryptic species, were subdivided by samples of each species (i.e. *P. elegans sensu stricto* and cryptic species) to conduct these analyses. Two-group AMOVA (*P. elegans sensu stricto* vs. cryptic species; Table 2b) maximized the  $F_{CT}$ , revealing that the 26.23% of the total variance was explained by differences between *P. elegans sensu stricto* and the

	Locality	Atlantic Ocean						Mediterranean Sea				
		Ré Island	Santoña	Ártabro Gulf	Cádiz	Tenerife	Lanzarote	Granada	Ebro Delta	Collioure	Marseille	Livorno
Atlantic Ocean	Ré Island	–	0.001*	0.175	0.210	0.000**	0.000**	0.096	0.000**	0.000**	0.000**	0.000**
	Santoña	0.022	–	0.135	0.029	0.000**	0.000**	0.550	0.000**	0.000**	0.000**	0.000**
	Ártabro Gulf	0.004	0.006	–	0.013	0.000**	0.000**	0.894	0.000**	0.000**	0.000**	0.000**
	Cádiz	0.004	0.011	0.009	–	0.000**	0.000**	0.314	0.000**	0.000**	0.000**	0.000**
	Tenerife	0.053	0.064	0.063	0.045	–	0.291	0.000**	0.000**	0.000**	0.000**	0.000**
	Lanzarote	0.053	0.066	0.063	0.041	0.003	–	0.000**	0.000**	0.000**	0.000**	0.000**
Mediterranean Sea	Granada	0.009	0.001	–0.005	0.004	0.038	0.034	–	0.000**	0.000**	0.000**	0.000**
	Ebro Delta	0.138	0.178	0.166	0.126	0.100	0.089	0.147	–	0.008	0.260	0.121
	Collioure	0.109	0.142	0.132	0.097	0.067	0.067	0.119	0.023	–	0.008	0.137
	Marseille	0.141	0.185	0.167	0.120	0.102	0.091	0.157	0.012	0.015	–	0.466
	Livorno	0.110	0.155	0.136	0.097	0.069	0.061	0.119	0.018	0.008	0.005	–

**Table 3.** Pairwise  $F_{ST}$  values (below the diagonal) between 11 sampling sites of *Palaemon elegans sensu stricto* (327 sampled individuals) as revealed by the microsatellite dataset (20 loci), and the corresponding  $P$ - values (above the diagonal). \*Significance after FDR correction ( $P < 0.0033$ ). \*\*Significance after Bonferroni correction ( $P < 0.00091$ ).

putative cryptic species. A minor but considerable proportion of the global variance (20.06%) was explained by differences among groups in three-group AMOVA (Atlantic localities and Granada of *P. elegans sensu stricto* vs. Mediterranean localities of *P. elegans sensu stricto* vs. localities of cryptic species; Table 2c). Nevertheless, a genetic structuring of the whole genotype dataset in two groups yielded a higher  $F_{CT}$  value.

Pairwise  $F_{ST}$  values were estimated between localities in which at least 10 individuals were sampled. Therefore, separating samples of Granada, Almería, Collioure, Livorno and Mallorca by species,  $F_{ST}$  values could not be calculated for some of these subdivided localities. Even so, pairwise  $F_{ST}$  values (Supplementary Table S3) suggested a marked and significant genetic differentiation in all the comparisons between localities of *P. elegans sensu stricto* and localities of cryptic species ( $F_{ST} \geq 0.190$ ), indicating the genetic divergence of these two species supported by the other analyses. Focussing on the population structure of *P. elegans sensu stricto*, pairwise  $F_{ST}$  values revealed significant genetic differentiation between the Atlantic and the Mediterranean localities (Table 3). Granada locality was connected with almost all the Atlantic localities and significantly differentiated from the remaining Mediterranean ones. Within the Mediterranean Sea, connectivity among Ebro Delta-Collioure-Marseille-Livorno was suggested by pairwise  $F_{ST}$  values. Along the Atlantic coast, connectivity was suggested among Ré Island-Santoña-Ártabro Gulf-Cádiz-Granada and some genetic differentiation was supported between each one of those and the Canarian localities. In other words, Tenerife and Lanzarote showed significant differentiation with all the localities regardless of the sampling source, while they were both connected to each other.

In order to increase the reliability of the study, the original microsatellite dataset (excluding locus Pe 11) was pared down to check the impact of the putative cryptic species on the genetic structure results of *P. elegans sensu stricto*. In fact, individuals belonging to the putative cryptic species presented a considerable amount of missing data in some loci, especially in loci Pe08 and Pe17. Thus, another STRUCTURE run was carried out removing these 69 individuals and the 4 individuals that could not be assigned to any cluster, so the microsatellite dataset exclusively contained genotypes of individuals considered as *P. elegans sensu stricto* (327 individuals). Delta $K$  statistic distribution peaked at  $K = 2$  (Supplementary Figure S1c), indicating again the partition of *P. elegans sensu stricto* samples into two clusters (Supplementary Figure S3a): one cluster joined Granada locality to the Atlantic localities (Ré Island, Santoña, Ártabro Gulf, Cádiz, Tenerife and Lanzarote), meanwhile the other cluster was formed by the Mediterranean localities (Almería, Ebro Delta, Collioure, Marseille, Livorno and Mallorca). The pattern of divergence between the sampling sites located on the Atlantic coast and those on the Mediterranean Sea was detected again for *P. elegans sensu stricto*. AMOVA was performed to test this two-group structure in this reduced dataset. When excluding the individuals of the putative cryptic species, differentiation among groups decreased ( $F_{CT} = 0.094$ ), explaining the 9.37% of the total variance (Supplementary Table S4).

By partitioning samples of *P. elegans sensu stricto* into two groups (Fig. 1b and Supplementary Figure S3a) according to the Bayesian clustering analysis, most of the individuals from Tenerife and Lanzarote were assigned to the Atlantic Ocean cluster as expected, but a considerable number of Canarian individuals could not be assigned to either of the two clusters (probability of membership to a cluster  $< 0.75$ ). Assessing the results of the STRUCTURE analysis using the *P. elegans sensu stricto* dataset, a peak of  $\Delta K$  and a plateau in the plot of  $K$  versus  $\ln \Pr(X|K)$  (Supplementary Figure S1d) were observed at  $K = 3$ , so the partition in three groups was plotted (Supplementary Figure S3b). With this grouping, the Mediterranean Sea cluster remained mainly invariable, whereas the Atlantic Ocean cluster was split into two subgroups, being then possible to clearly assign individuals from the Canary Islands localities to one of these two Atlantic clusters. Three-group AMOVA using this dataset unveiled that the 8.80% of the global variance is due to differences among groups, so the  $F_{CT}$  value ( $F_{CT} = 0.088$ ) was lesser respecting to two-group AMOVA (Supplementary Table S4). Overall, the most likely genetic structure of *P. elegans sensu stricto* is two populations, one of them including the Atlantic localities and Granada, and the other encompassing the Mediterranean localities. Nevertheless, it is noteworthy to highlight that there is a substructure within the Atlantic population, with a subpopulation not exclusively but mainly formed by individuals from

the Canary Islands. This substructure is in line with the genetic differentiation between Tenerife and Lanzarote and the remaining Atlantic localities suggested by the *FST* pairwise values. Finally, isolation-by-distance (IBD) was retrieved in *P. elegans sensu stricto* by the Mantel test, as genetic and geographical distances resulted to be significantly correlated ( $R=0.37$ ,  $P=0.001$ ). Another two Mantel tests were performed to test IBD within each population of *P. elegans sensu stricto*. No IBD was detected within the Mediterranean basin ( $R=0.00$ ,  $P=0.750$ ), meanwhile a significant relationship between genetic and geographical distances was found along the Atlantic population ( $R=0.40$ ,  $P=0.015$ ).

## Discussion

In this study, microsatellite markers were used for the analysis of the genetic diversity and population structure of the shrimp *P. elegans*. Levels of polymorphism were in line with those found in the counterpart common littoral shrimp *P. serratus*<sup>25,29</sup>. Sampling sites showed similar levels of genetic diversity (*Na*, *A*, *AR*, *Ho* or *He*). Thirty-one locality-locus combinations significantly deviated from HWE, mainly caused by heterozygote deficiency (29 out of the 31 cases) reflected in positive *FIS* values. Stutter-related scoring errors were discarded as the source of HWE departures because markers with ambiguous interpretation patterns were not included in the microsatellite set during the characterization work<sup>27</sup> nor reported during the genotyping step here. Different causes for HWE deviations are frequently invoked such as inbreeding, Wahlund effect or null alleles. Inbreeding should equally affect all loci, generating similar heterozygote deficit across them. Despite global positive *FIS* values were obtained in all localities, great differences in *FIS* values across loci were observed, discarding inbreeding. Heterozygote deficit caused by subpopulation structure (Wahlund effect) could not be completely ruled out in localities where an arbitrary admixture of individuals from putative different species was found. Nevertheless, the presence of null alleles is the most likely explanation for the observed departures from HWE. The instability of the flanking sequences of microsatellites leads to some alleles could not be amplified and consequently dropped out, resulting in a homozygote excess. Although microsatellite null alleles are widespread, marine invertebrates have demonstrated particularly higher frequencies than other groups<sup>30,31</sup>. The 12 loci that deviated from HWE due to heterozygote deficit showed evidence of occurrence of null alleles, even though mainly in a low proportion. Only one locus, the locus Pe11, showed certain consistency in high null allele frequency estimates across localities and overall. For this reason, we decided to exclude locus Pe11 in population structure analyses as it was advised in González-Castellano et al.<sup>27</sup>.

Microsatellite data provided evidence for the existence of two clearly genetically distinct groups within *P. elegans*. Some individuals from Mediterranean localities showed a marked genetic divergence from the remaining specimens without following any geographical or oceanographic pattern. Since all shrimps from Granada, Almería, Collioure, Livorno and Mallorca were captured simultaneously in the same rockpools and unambiguously morphologically identified as *P. elegans* afterwards, we concluded that the explanation for the divergent individuals is the presence of a cryptic and sympatric species existing within *P. elegans*. Previous studies using mtDNA markers reported that some individuals from the Mediterranean Sea exhibited divergent haplotypes respecting the rest of the sampled collection, being those named *P. elegans* type III and considered as a cryptic species<sup>8,12,13</sup>. In fact, the individuals assigned with our microsatellite set to a cryptic species were confirmed to correspond with those *P. elegans* type III by their mtDNA COI sequences. The alternative hypothesis to explain this divergence in nuclear markers among our samples would be strong genetic structuring among individuals who live sympatrically and without exhibiting evident intrinsic reproductive barriers, what seems extremely unlikely.

Cryptic species are distributed among taxa, but they are particularly prone in the marine environment<sup>32</sup>, being detected within various species using microsatellite markers<sup>33–35</sup>. Here, the 400 sampled shrimps were assigned to one of the two unveiled genetic groups, i.e. *P. elegans sensu stricto* or cryptic species. Only two individuals from Granada and two from Mallorca could not be assigned to one specific cluster, most likely because they presented missing data at several loci. Consequently, no hybrids between *P. elegans sensu stricto* and the cryptic species were found. Hence, unknown reproductive isolating mechanisms between these two species should be acting to explain the absence of genetic hybrids. The main question to be solved is the origin of the cryptic species. Given that it was confined in the Mediterranean Sea, it should be originated there. The Mediterranean basin has a very complex geological history that largely contributed to its high level of biodiversity and endemism<sup>15</sup>. Reuschel et al.<sup>8</sup> postulated that *P. elegans* type III specimens might have their origin during the Messinian Salinity Crisis in late Miocene (around 5.6 Myr ago), when the Strait of Gibraltar was closed, isolating the Mediterranean Sea from the Atlantic Ocean and determining the extensive desiccation of the Mediterranean basin<sup>36</sup>. According to Reuschel et al.<sup>8</sup>, isolated *P. elegans* type III ancestors had to survive in local refugia in the Mediterranean Sea and genetically diverge from the Atlantic population of *P. elegans*, giving rise to the cryptic species detected within *P. elegans* nowadays. With the re-opening of the Strait of Gibraltar, the Mediterranean basin was flooded again with Atlantic water, so Atlantic *P. elegans* individuals should be re-introduced, but they should no longer be able to reproduce with the cryptic species ancestors, since none hybrids are currently detected. Similarly, the Messinian Salinity Crisis was considered the origin of two sister crab species of the genus *Carcinus*<sup>37</sup> and of two sister brittle star species of the genus *Ophiothrix*<sup>38</sup>. Our analyses using microsatellite loci support the existence of two species within *P. elegans*, providing nuclear evidence for the divergence signatures previously detected at mitochondrial level. Future efforts should be directed to clarify the origin of the cryptic species by phylogenetic or phylogenomic studies with precise calibration and to revise the taxonomy of *P. elegans* complex.

Regarding the population biology of *P. elegans sensu stricto* (*P. elegans* hereafter), a genetic structure was revealed in the whole study. Genetic differentiation analyses (Bayesian analysis, AMOVA and *FST* statistics) showed a pattern of two main populations: one composed by the Atlantic localities and Granada, and another composed by the Mediterranean localities. In other words, a clear geographic distinction between Atlantic and

Mediterranean sampling sites was detected, with an extension of the Atlantic genotypes into Mediterranean Sea waters. Granada locality was genetically closer to the Atlantic localities than to the other Mediterranean localities. Therefore, a biogeographical barrier restricting the gene flow between these two populations should be located between Granada (Atlantic population) and Almería (Mediterranean population). Precisely, this genetic break matched with the location of the Almería-Orán Front, a semi-permanent dynamic oceanographic front consisted in two anticyclonic gyres that entails abrupt changes of temperature and salinity<sup>39</sup>. The Almería-Orán Front has been considered as the most important hydrographic boundary to gene flow between the Atlantic Ocean and Mediterranean Sea surface waters<sup>11</sup>. Indeed, it was confirmed that the Almería-Orán Front acts on the population structure of several species as in *P. elegans*, e.g. the cuttlefish *Sepia officinalis*<sup>40</sup>, some fish species<sup>41</sup>, the barnacle *Chthamalus montagu*<sup>42</sup> or the wedge clam *Donax trunculus*<sup>43</sup> among others (see review in Patarnello et al.<sup>11</sup>). Notwithstanding, it is known that population connectivity across the Atlantic–Mediterranean transition is reduced in other species by different oceanographic discontinuities, namely the Strait of Gibraltar<sup>24,44</sup>, the Balearic Front and the Ibiza Channel<sup>45,46</sup>. Most recently, in the congeneric common littoral shrimp *Palaemon serratus*, both mtDNA<sup>47</sup> and nuclear<sup>25</sup> markers have reported strong genetic differentiation among Atlantic and Mediterranean localities with an unusual phylogeographical break located west of the Strait of Gibraltar in the Gulf of Cádiz. In the present study, microsatellite data unravelled that the Almería-Orán Front was the only oceanographic discontinuity affecting the population structure of *P. elegans* across the Atlantic–Mediterranean transition, so our results agreed with the phylogeographic pattern described in Reuschel et al.<sup>8</sup> using mitochondrial data. However, it is important to note that the Almería-Orán Front experiences seasonal fluctuations in intensity, size and position<sup>39</sup> that might mitigate its effects as oceanographic barrier between the populations of *P. elegans*.

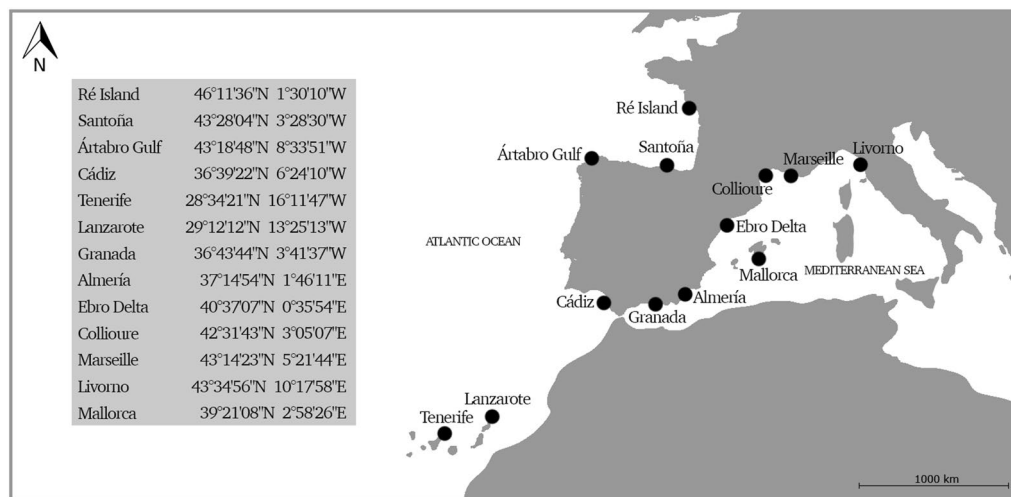
Differentiation was more complex within each group. Within the Mediterranean population, all localities showed no genetic differentiation among them in the light of our data. Our shrimp samples only covered the northern of the Western Mediterranean basin, so it would be interesting to analyse more Mediterranean localities to study population connectivity throughout the entire Mediterranean Sea. As for the Atlantic group, genetic analyses indicated some substructure within this population. Specifically, Canary Islands localities (Tenerife and Lanzarote) showed low but significant genetic differentiation from the remaining Atlantic localities (Ré Island, Santoña, Ártabro Gulf, Cádiz and Granada) as suggested by the pairwise *F<sub>ST</sub>* values and also observed in the Bayesian analysis. Thus, the above results allowed us to report two subgroups within the Atlantic population: one formed by the South of France and the Iberian Peninsula localities; and another group comprising the Canary Islands localities. The lack of connectivity between these subgroups is consistent with the isolation-by-distance pattern observed along the Atlantic population of *P. elegans*. In this regard, lack of gene flow between Macaronesian Islands (Azores, Madeira, Canary Islands and Cape Verde) and continental locations has been also reported for various marine invertebrates, such as limpets of the genus *Patella*<sup>48</sup>, starfishes of the genus *Marthasterias* spp.<sup>49,50</sup> or the sea cucumber *Holothuria* (*Holothuria*) *mammata*<sup>51</sup>, as well as for fishes as the triplefin *Tripterygion delaisi*<sup>52</sup> and the white seabream *Diplodus sargus*<sup>53</sup>. The remoteness of the Canary Islands from the Atlantic mainland sampling sites may explain the weak differentiation of Canarian *P. elegans* samples. The duration of the larval stage compared to the great distance that separates the Canary Islands and the continent may not be enough to prevent this differentiation. Related to this, present-day surface currents in Macaronesia, mainly the Azores Current and the Canary Current<sup>54</sup>, could also favour the relative isolation of the Canary Islands by disrupting larval dispersion. A more extensive sampling would be needed to reveal whether there is more genetic structuring along the Atlantic distribution of this shrimp.

Finally, the microsatellite loci used in this study were isolated and characterized from Atlantic individuals<sup>27</sup>, so they are specific markers for *P. elegans sensu stricto*. In order to study the genetic structure of the cryptic species found in the Mediterranean Sea, firstly we recommend revising its status as an independent new species, and then to develop novel specific microsatellite markers to carry out population analyses with them. De novo development of species-specific microsatellites is strongly recommended. Cross-amplification from congeneric species is not generally feasible because inherent problems like allele size homoplasy, polymorphisms biases, null alleles presence, broken repeats motifs or amplification of non-orthologous loci could arise<sup>55–58</sup>.

In conclusion, this study encompasses the first assessment of the genetic diversity and population structure of the shrimp *P. elegans* using microsatellite markers. Two clearly genetically distinct groups were detected within what is taxonomically considered as *P. elegans*, indicating the existence of a Mediterranean cryptic species within *P. elegans*. Furthermore, *P. elegans sensu stricto* is genetically structured in two populations: the Atlantic population and the Mediterranean population. The genetic break that separates these populations was found across the Atlantic–Mediterranean transition, matching with the location of the Almería-Orán Front. The microsatellite loci were sensitive enough to also reveal certain differentiation within the Atlantic population mainly involving individuals from the Canary Islands, which may be explained by an isolation-by-distance pattern.

## Material and methods

Specimens of *P. elegans* used in this study were collected inshore between 2012 and 2019 from 13 European sampling sites: Ré Island (France), Santoña (Spain), Ártabro Gulf (Spain), Cádiz (Spain), Tenerife (Spain), Lanzarote (Spain), Granada (Spain), Almería (Spain), Ebro Delta (Spain), Collioure (France), Marseille (France), Livorno (Italy) and Mallorca (Spain) (Fig. 2). Animals were captured in intertidal rockpools with a fish trap and ethanol-stored at laboratory. Morphological identification of the 400 sampled shrimps was thoroughly carried out according to González-Ortegón & Cuesta<sup>1</sup>. Genomic DNA was extracted from 25 mg of abdominal muscle tissue using the NZY Tissue gDNA Isolation kit (NZYTech) following the manufacturer's instructions. Extracted DNA quality and concentration were determined with the NanoDrop ND-1000 spectrophotometer (Thermo Fisher Scientific).



**Figure 2.** Geographical location and coordinates of the sampling sites of *Palaemon elegans*. Map was newly generated by the authors using AutoCAD software.

Individual genotypes were obtained for 21 polymorphic microsatellite loci (Pe01–Pe21: MH078079–MH078099) previously developed by González-Castellano et al.<sup>27</sup> for *P. elegans* and following the wet-lab protocol described there. Fluorescently labelled polymerase chain reaction (PCR) products were run on a 3130xl Genetic Analysis System (Applied Biosystems) for fragment analysis in the Scientific Research Support Services (University of A Coruña), using GeneScan 500 (–250) ROX internal size standard (Applied Biosystems). Geneious 8.0.5<sup>59</sup> was used for fluorescent profiles analysing and allele peaks calling.

Genetic variability was assessed with standard descriptive statistics. Number of alleles ( $N_a$ ), mean number of alleles ( $A$ ), percentage of polymorphism ( $P$ ), observed heterozygosity ( $H_o$ ), expected heterozygosity ( $H_e$ ) and deviations from Hardy–Weinberg equilibrium ( $HWE$ ) for each locus and/or locality were estimated using GenALEX 6.5<sup>60</sup>. Significance levels were adjusted applying the sequential Bonferroni correction whenever multiple tests were performed. Allelic richness per locality was calculated using FSTAT 2.9.4<sup>61</sup>. GENEPOP 4.2<sup>62</sup> was used to estimate inbreeding coefficient ( $FIS$ ) and to test linkage disequilibrium between pairs of loci using Fisher's exact test. Significance was computed by a Markov chain method using 10,000 dememorizations, 5,000 batches and 5,000 iterations per batch. Departures from selective neutrality were evaluated for each locus applying the Ewens–Watterson test using the algorithm given in Manly<sup>63</sup> and implemented in POPGENE 1.32<sup>64</sup> after 1,000 simulations. Null allele frequency was estimated using FreeNA<sup>28</sup> and the EM algorithm<sup>65</sup> with a number of bootstrap replicates fixed to 10,000.

Genetic structure was assessed using the Bayesian model-based clustering method implemented in STRUCTURE 2.3.4<sup>66</sup>. The admixture model with correlated allele frequencies between populations<sup>67</sup> was used. A burn-in of 100,000 iterations followed by 300,000 iterations for parameter estimation was set up. Each simulation was run 10 times, exploring values for  $K$  ranging from 1 to  $N + 2$ , where  $N$  was the number of localities. Results were processed using STRUCTURE HARVESTER<sup>68</sup>. The best partition of the data was determined by examining both the log probability of the data ( $\ln \Pr(X|K)$ ) and the  $\Delta K$  statistic following Evanno, Regnaut, & Goudet<sup>69</sup>. CLUMPP 1.1.2<sup>70</sup> was used to permute the admixture coefficients for the several independent runs resulting for the chosen  $K$ -value. STRUCTURE PLOT 2.0<sup>71</sup> was employed to visualize the output from CLUMPP. Genetic discrimination among samples was additionally assessed through a factorial correspondence analysis (FCA) using GENETIX 4.05<sup>72</sup>.

Genetic differentiation was estimated as pairwise  $FST$  values (10,000 permutations) between localities using ARLEQUIN 3.5.1.2<sup>73</sup>. Significance level was adjusted applying sequential Bonferroni correction for multiple comparisons and false discovery rate (FDR) correction (B–Y method as described in Narum<sup>74</sup>). Several hierarchical locus-by-locus analyses of molecular variance (AMOVA) were carried out using ARLEQUIN 3.5.1.2<sup>73</sup> in order to estimate partitioning of genetic variance among groups ( $FCT$ ), among localities within groups ( $FSC$ ) and within localities ( $FST$ ). Individuals were pooled according to a geographic/oceanographic criterion or based on the results previously obtained with STRUCTURE to test different groupings.

Lastly, the relationship between genetic distance and geographic distance was assessed with a Mantel test<sup>75</sup> to test isolation-by-distance (IBD) using the program ISOLDE implemented in GENEPOP 4.2<sup>62</sup> with 10,000 permutations. Pairwise genetic differentiation among localities was estimated using linearized  $FST$  [ $FST/(1 - FST)$ ] to construct a genetic distance matrix. Pairwise geographic distance between localities was measured in kilometres following the coastline in Google Earth (Google Inc.) to create a geographic distance matrix.

### Data availability

Accession codes: Sequences containing the microsatellite loci were deposited in GenBank under Accession Numbers MH078079–MH078099. Genotype dataset generated during the current study is available from the corresponding authors on reasonable request.



Received: 22 January 2020; Accepted: 15 June 2020

Published online: 01 July 2020

## References

- González-Ortegón, E. & Cuesta, J. A. An illustrated key to species of *Palaemon* and *Palaemonetes* (Crustacea: Decapoda: Caridea) from European waters, including the alien species *Palaemon macrodactylus*. *J. Mar. Biol. Assoc. UK* **86**, 93–102. <https://doi.org/10.1017/S0025315406012896> (2006).
- Taylor, A. C. & Spicer, J. I. Metabolic responses of the prawns *Palaemon elegans* and *P. serratus* (Crustacea: Decapoda) to acute hypoxia and anoxia. *Mar. Biol.* **95**, 521–530. <https://doi.org/10.1007/BF00393095> (1987).
- Janas, U., Pilka, M. & Lipinska, D. Temperature and salinity requirements of *Palaemon adspersus* Rathke 1837 and *Palaemon elegans* Rathke, 1837. Do they explain the occurrence and expansion of prawns in the Baltic Sea?. *Mar. Biol. Res.* **9**, 293–300. <https://doi.org/10.1080/17451000.2012.739699> (2013).
- d'Acoz, U. Inventaire et distribution des crustacés décapodes de l'Atlantique nord-oriental, de la Méditerranée et des eaux continentales adjacentes au nord de 25°N. *Collection Patrimoines Naturels (Muséum National d'Histoire Naturelle (S.P.N.))* [https://doi.org/10.1016/s0990-7440\(02\)01163-4](https://doi.org/10.1016/s0990-7440(02)01163-4) (1999).
- Grabowski, M. Rapid colonization of the Polish Baltic coast by an Atlantic palaemonid shrimp *Palaemon elegans* Rathke, 1837. *Aquat. Invasions* **1**, 116–123. <https://doi.org/10.3391/ai.2006.1.3.3> (2006).
- Katajisto, T., Kotta, J., Lehtiniemi, M., Malavin, S. A. & Panov, V. E. *Palaemon elegans* Rathke, 1837 (Caridea: Palaemonoidea: Palaemonidae) established in the Gulf of Finland. *Bioinvasions Rec.* **2**, 125–132. <https://doi.org/10.3391/bir.2013.2.2.05> (2013).
- González-Ortegón, E., Sargent, P., Pohle, G. & Martínez-Lage, A. The Baltic prawn *Palaemon adspersus* Rathke, 1837 (Decapoda, Caridea, Palaemonidae): First record, possible establishment, and illustrated key of the subfamily Palaemoninae in northwest Atlantic waters. *Aquat. Invasions* **10**, 299–312. <https://doi.org/10.3391/ai.2015.10.3.05> (2015).
- Reuschel, S., Cuesta, J. A. & Schubart, C. D. Marine biogeographic boundaries and human introduction along the European coast revealed by phylogeography of the prawn *Palaemon elegans*. *Mol. Phylogenet. Evol.* **55**, 765–775. <https://doi.org/10.1016/j.ympev.2010.03.021> (2010).
- Ward, R. D. The importance of identifying spatial population structure in restocking and stock enhancement programmes. *Fish. Res.* **80**, 9–18. <https://doi.org/10.1016/j.fishres.2006.03.009> (2006).
- Pascual, M., Rives, B., Schunter, C. & Macpherson, E. Impact of life history traits on gene flow: a multispecies systematic review across oceanographic barriers in the Mediterranean Sea. *PLoS ONE* **12**, e0176419. <https://doi.org/10.1371/journal.pone.0176419> (2017).
- Patarnello, T., Volckaert, F. A. M. J. & Castilho, R. Pillars of Hercules: is the Atlantic–Mediterranean transition a phylogeographical break?. *Mol. Ecol.* **16**, 4426–4444. <https://doi.org/10.1111/j.1365-294X.2007.03477.x> (2007).
- Bilgin, R., Utkan, M. A., Kalkan, E., Karhan, S. & Bekbolet, M. DNA barcoding of twelve shrimp species (Crustacea: Decapoda) from Turkish seas reveals cryptic diversity. *Mediterr. Mar. Sci.* **16**, 36–45. <https://doi.org/10.12681/mms.548> (2014).
- Deli, T., Pfaller, M. & Schubart, C. D. Phylogeography of the littoral prawn species *Palaemon elegans* (Crustacea: Caridea: Palaemonidae) across the Mediterranean Sea unveils disparate patterns of population genetic structure and demographic history in the two sympatric genetic types II and III. *Mar. Biodiv.* <https://doi.org/10.1007/s12526-017-0711-6> (2018).
- Bickford, D. *et al.* Cryptic species as a window on diversity and conservation. *Trends Ecol. Evol.* **22**, 148–155. <https://doi.org/10.1016/j.tree.2006.11.004> (2007).
- Lejeune, C., Chevaldonné, P., Pergent-Martini, C., Boudouresque, C. F. & Pérez, T. Climate change effects on a miniature ocean: the highly diverse, highly impacted Mediterranean Sea. *Trends Ecol. Evol.* **25**, 250–260. <https://doi.org/10.1016/j.tree.2009.10.009> (2010).
- Fernández, M. V., Heras, S., Maltagliati, F. & Roldán, M. I. Deep genetic divergence in giant red shrimp *Aristaeomorpha foliacea* (Risso, 1827) across a wide distributional range. *J. Sea Res.* **76**, 146–153. <https://doi.org/10.1016/j.seares.2012.08.004> (2013).
- Fernández, M. V., Heras, S., Vinas, J., Maltagliati, F. & Roldán, M. I. Multilocus comparative phylogeography of two Aristeid shrimps of high commercial interest (*Aristeus antennatus* and *Aristaeomorpha foliacea*) reveals different responses to past environmental changes. *PLoS ONE* **8**, e59033. <https://doi.org/10.1371/journal.pone.0059033> (2013).
- Barba, De. *et al.* High-throughput microsatellite genotyping in ecology: improved accuracy, efficiency, standardization and success with low-quantity and degraded DNA. *Mol. Ecol. Resour.* **17**, 492–507. <https://doi.org/10.1111/1755-0998.12594> (2017).
- Goldstein, D. B. & Schlotterer, C. *Microsatellites—Evolution and Applications*. Oxford University Press, Oxford. <https://doi.org/10.1017/S0016672300239141> (1999).
- Wright, J. M. & Bentzen, P. Microsatellites: genetic markers for the future. *Mol. Genet. Fish.* **1**, 117–121. [https://doi.org/10.1007/978-94-011-1218-5\\_7](https://doi.org/10.1007/978-94-011-1218-5_7) (1995).
- Valles-Jiménez, R., Cruz, P. & Pérez-Enríquez, R. Population genetic structure of pacific white shrimp (*Litopenaus vannamei*) from Mexico to Panama: microsatellite DNA variation. *Mar. Biotechnol.* **6**, 475–484. <https://doi.org/10.1007/s10126-004-3138-6> (2004).
- Waqairatu, S. S. *et al.* Genetic analysis of black tiger shrimp (*Penaeus monodon*) across its natural distribution range reveals more recent colonization of Fiji and other South Pacific islands. *Ecol. Evol.* **2**, 2057–2071. <https://doi.org/10.1002/ece3.316> (2012).
- Jorde, P. E. *et al.* Genetically distinct populations of northern shrimp, *Pandalus borealis*, in the North Atlantic: adaptation to different temperatures as an isolation factor. *Mol. Ecol.* **24**, 1742–1757. <https://doi.org/10.1111/mec.13158> (2015).
- Heras, S., Planella, L., García-Marín, J. L., Vera, M. & Roldán, M. I. Genetic structure and population connectivity of the blue and red shrimp *Aristeus antennatus*. *Sci. Rep.* **9**, 1–11. <https://doi.org/10.1038/s41598-019-49958-5> (2019).
- Perina, A. *et al.* Assessment of genetic diversity and population structure of the common littoral shrimp *Palaemon serratus* (Pennant, 1777) by microsatellites: towards a sustainable management. *Aquatic Conserv. Mar. Freshw. Ecosyst.* <https://doi.org/10.1002/aqc.3011> (2019).
- Davey, J. W. *et al.* Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nat. Rev. Genet.* **12**, 499–510. <https://doi.org/10.1038/nrg3012> (2011).
- González-Castellano, I., Perina, A., González-Tizón, A. M., Torrecilla, Z. & Martínez-Lage, A. Isolation and characterization of 21 polymorphic microsatellite loci for the rockpool shrimp *Palaemon elegans* using Illumina MiSeq sequencing. *Sci. Rep.* **8**, 17197. <https://doi.org/10.1038/s41598-018-35408-1> (2018).
- Chapuis, M. P. & Estoup, A. Microsatellite null alleles and estimation of population differentiation. *Mol. Biol. Evol.* **24**, 621–631. <https://doi.org/10.1093/molbev/msl191> (2007).
- Perina, A., González-Tizón, A. M., Vizcaíno, A., González-Ortegón, E. & Martínez-Lage, A. Isolation and characterization of 20 polymorphic microsatellite loci in *Palaemon serratus* and cross-amplification in *Palaemon* species by 454 pyrosequencing. *Conserv. Genet. Resour.* **8**, 169–196. <https://doi.org/10.1007/s12686-016-0549-4> (2016).
- Dailianis, T., Tsigenopoulos, C. S., Dounas, C. & Voultsiadou, E. Genetic diversity of the imperilled bath sponge *Spongia officinalis* Linnaeus, 1759 across the Mediterranean Sea: patterns of population differentiation and implications for taxonomy and conservation. *Mol. Ecol.* **20**, 3757–3772. <https://doi.org/10.1111/j.1365-294X.2011.05222.x> (2011).
- Plough, L. V. & Marko, P. B. Characterization of microsatellite loci and repeat density in the gooseneck barnacle, *Pollicipes elegans*, using next generation sequencing. *J. Hered.* **105**, 136–142. <https://doi.org/10.1093/jhered/est064> (2014).

32. Boissin, E., Féral, J. P. & Chenuil, A. Defining reproductively isolated units in a cryptic and syntopic species complex using mitochondrial and nuclear markers: the brooding brittle star, *Amphipholis squamata* (Ophiuroidea). *Mol. Ecol.* **17**, 1732–1744. <https://doi.org/10.1111/j.1365-294X.2007.03652.x> (2008).
33. Griffiths, A. M. *et al.* Molecular markers reveal spatially segregated cryptic species in a critically endangered fish, the common skate (*Dipturus batis*). *Proc. R. Soc. B. Biol. Sci.* **277**, 1497–1503. <https://doi.org/10.1098/rspb.2009.2111> (2010).
34. Andrews, K. R., Norton, E. L., Fernandez-Silva, I., Portner, E. & Goetze, E. Multilocus evidence for globally distributed cryptic species and distinct populations across ocean gyres in a mesopelagic copepod. *Mol. Ecol.* **23**, 5462–5479. <https://doi.org/10.1111/mec.12950> (2014).
35. Naughton, K. M., O'Hara, T. D., Appleton, B. & Gardner, M. G. Sympatric cryptic species in the crinoid genus *Cenolia* (Echinodermata: Crinoidea: Comasteridae) delineated by sequence and microsatellite markers. *Mol. Phylogenet. Evol.* **78**, 160–171. <https://doi.org/10.1016/j.ympev.2014.05.006> (2014).
36. García-Castellanos, D. *et al.* Catastrophic flood of the Mediterranean after the Messinian salinity crisis. *Nature* **462**, 778. <https://doi.org/10.1038/nature08555> (2009).
37. Marino, I. A., Pujolar, J. M. & Zane, L. Reconciling deep calibration and demographic history: Bayesian inference of post glacial colonization patterns in *Carcinus aestuarii* (Nardo, 1847) and *C. maenas* (Linnaeus, 1758). *PLOS ONE* **6**, e28567. <https://doi.org/10.1371/journal.pone.0028567> (2011).
38. Pérez-Portela, R., Almada, V. & Turon, X. Cryptic speciation and genetic structure of widely distributed brittle stars (Ophiuroidea) in Europe. *Zool. Scr.* **42**, 151–169. <https://doi.org/10.1111/j.1463-6409.2012.00573.x> (2013).
39. Tintoré, J., La Violette, P. E., Blade, I. & Cruzado, A. A study of an intense density front in the eastern Alboran Sea: the Almeria-Oran front. *J. Phys. Oceanogr.* **18**, 1384–1397. [https://doi.org/10.1175/1520-0485\(1988\)018<1384:ASOAIID>2.0.CO;2](https://doi.org/10.1175/1520-0485(1988)018<1384:ASOAIID>2.0.CO;2) (1988).
40. Pérez-Losada, M., Guerra, A., Carvalho, G. R., Sanjuan, A. & Shaw, P. W. Extensive population subdivision of the cuttlefish *Sepia officinalis* (Mollusca: Cephalopoda) around the Iberian Peninsula indicated by microsatellite DNA variation. *Heredity* **89**, 417–424. <https://doi.org/10.1038/sj.hdy.6800160> (2002).
41. Galarza, J. A. *et al.* The influence of oceanographic fronts and early-life-history traits on connectivity among littoral fish species. *Proc. Natl. Acad. Sci. USA* **106**, 1473–1478. <https://doi.org/10.1073/pnas.0806804106> (2009).
42. Pannacciulli, F. G., Maltagliati, F., De Guttry, C. & Achituv, Y. Phylogeography on the rocks: the contribution of current and historical factors in shaping the genetic structure of *Chthamalus montagui* (Crustacea, Cirripedia). *PLoS ONE* **12**, e0178287. <https://doi.org/10.1371/journal.pone.0178287> (2017).
43. Nantón, A. *et al.* Microsatellite variation in *Donax trunculus* from the Iberian Peninsula, with particular attention to Galician estuaries (NW Spain). *Estuar. Coast. Shelf Sci.* **197**, 27–34. <https://doi.org/10.1016/j.ecss.2017.08.011> (2017).
44. Fratini, S. *et al.* Unravelling population genetic structure with mitochondrial DNA in a notional panmictic coastal crab species: sample size makes the difference. *BMC Evol. Biol.* **16**, 150. <https://doi.org/10.1186/s12862-016-0720-2> (2016).
45. Schunter, C. *et al.* Matching genetics with oceanography: directional gene flow in a Mediterranean fish species. *Mol. Ecol.* **20**, 5167–5181. <https://doi.org/10.1111/j.1365-294X.2011.05355.x> (2011).
46. García-Merchán, V. H. *et al.* Phylogeographic patterns of decapod crustaceans at the Atlantic–Mediterranean transition. *Mol. Phylogenet. Evol.* **62**, 664–672. <https://doi.org/10.1016/j.ympev.2011.11.009> (2012).
47. Weiss, R. *et al.* Genetic differentiation between Mediterranean and Atlantic populations of the common prawn *Palaemon serratus* (Crustacea: Palaemonidae) reveals uncommon phylogeographic break. *J. Mar. Biol. Assoc. UK* **98**, 1425–1434. <https://doi.org/10.1017/S0025315417000492> (2018).
48. Sá-Pinto, A., Branco, M., Sayanda, D. & Alexandrino, P. Patterns of colonization, evolution and gene flow in species of the genus *Patella* in the Macaronesian Islands. *Mol. Ecol.* **17**, 519–532. <https://doi.org/10.1111/j.1365-294X.2007.03563.x> (2008).
49. Pérez-Portela, R., Villamor, A. & Almada, V. Phylogeography of the sea star *Marthasterias glacialis* (Asteroidea, Echinodermata): deep genetic divergence between mitochondrial lineages in the north-western mediterranean. *Mar. Biol.* **157**, 2015–2028. <https://doi.org/10.1007/s00227-010-1470-0> (2010).
50. Pérez-Portela, R., Rius, M. & Villamor, A. Lineage splitting, secondary contacts and genetic admixture of a widely distributed marine invertebrate. *J. Biogeogr.* **44**, 446–460. <https://doi.org/10.1111/jbi.12917> (2017).
51. Borrero-Pérez, G. H., González-Wangüemert, M., Marcos, C. & Pérez-Ruzafa, A. Phylogeography of the Atlanto-Mediterranean Sea cucumber *Holothuria (Holothuria) mammata*: the combined effects of historical processes and current oceanographical pattern. *Mol. Ecol.* **20**, 1964–1975. <https://doi.org/10.1111/j.1365-294X.2011.05068.x> (2011).
52. Domingues, V. S., Almada, V. C., Santos, R. S., Brito, A. & Bernardi, G. Phylogeography and evolution of the triplefin *Tripterygion delaisi* (Pisces, Blennioidei). *Mar. Biol.* **150**, 509–519. <https://doi.org/10.1007/s00227-006-0367-4> (2007).
53. González-Wangüemert, M., Froufe, E., Pérez-Ruzafa, A. & Alexandrino, P. Phylogeographical history of the white seabream *Diplodus sargus* (Sparidae): implications for insularity. *Mar. Biol. Res.* **7**, 250–260. <https://doi.org/10.1080/17451000.2010.499438> (2011).
54. Mason, E., Coombs, S. & Oliveira, P.B. An overview of the literature concerning the oceanography of the eastern North Atlantic region. *Relatórios Científicos e Técnicos IPIMAR Série digital no. 33*.
55. Primmer, R. C., Painter, N. J., Koskinen, T. M., Palo, U. J. & Merilä, J. Factors affecting avian cross-species microsatellite amplification. *J. Avian Biol.* **36**, 348–360. <https://doi.org/10.1111/j.0908-8857.2005.03465.x> (2005).
56. Barbará, T. *et al.* Cross-species transfer of nuclear microsatellite markers: potential and limitations. *Mol. Ecol.* **16**, 3759–3767. <https://doi.org/10.1111/j.1365-294X.2007.03439.x> (2007).
57. Rutkowski, R., Sielezniew, M. & Szostak, A. Contrasting levels of polymorphism in cross-amplified microsatellites in two endangered xerothermophilous, obligatorily myrmecophilous, butterflies of the genus *Phengaris* (Maculinea) (Lepidoptera: Lycaenidae). *Eur. J. Entomol.* **106**, 457–469. <https://doi.org/10.14411/eje.2009.058> (2009).
58. Yue, G. H., Balazs, K. & Laszlo, O. A new problem with cross-species amplification of microsatellites: generation of non-homologous products. *Zool. Res.* **31**, 131–140. <https://doi.org/10.3724/SP.J.1141.2010.02131> (2010).
59. Kears, M. *et al.* Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28**, 1647–1649. <https://doi.org/10.1093/bioinformatics/bts199> (2012).
60. Peakall, R. & Smouse, P. E. GenAlEx 6: genetic analysis in Excel. Population genetic software for teaching and research. *Mol. Ecol. Resour.* **6**, 288–295. <https://doi.org/10.1111/j.1471-8286.2005.01155.x> (2006).
61. Goudet, J. FSTAT version 2.9.4: a program to estimate and test population genetics parameters. Updated from Goudet [1995]. Available from <https://www.unil.ch/izea/software/fstat.html> (2003).
62. Rousset, F. Genepop'007: a complete re-implementation of the Genepop software for Windows and Linux. *Mol. Ecol. Resour.* **8**, 103–106. <https://doi.org/10.1111/j.1471-8286.2007.01931.x> (2008).
63. Manly, B. F. J. *The statistics of natural selection* (Chapman and Hall, London, 1985).
64. Yeh, F. C. *et al.* Popgene, the User Friendly Shareware for Population Genetic Analysis (Molecular Biology and Biotechnology Centre, University of Alberta, Alberta, Canada, 1997).
65. Dempster, A. P., Laird, N. M. & Rubin, D. B. Maximum likelihood from incomplete data via the EM algorithm. *J. R. Stat. Soc. Ser. B (methodological)* <https://doi.org/10.1111/j.2517-6161.1977.tb01600.x> (1977).
66. Pritchard, J. K., Stephens, M. & Donnelly, P. Inference of population structure using multilocus genotype data. *Genetics* **155**, 945–959 (2000).

67. Hubisz, M. J., Falush, D., Stephens, M. & Pritchard, J. K. Inferring weak population structure with the assistance of sample group information. *Mol. Ecol. Resour.* **9**, 1322–1332. <https://doi.org/10.1111/j.1755-0998.2009.02591.x> (2009).
68. Earl, D. A. & von Holdt, B. M. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv Genet. Resour.* **4**, 359–361. <https://doi.org/10.1007/s12686-011-9548-7> (2012).
69. Evanno, G., Regnaut, S. & Goudet, J. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol. Ecol.* **14**, 2611–2620. <https://doi.org/10.1111/j.1365-294X.2005.02553.x> (2005).
70. Jakobsson, M. & Rosenberg, N. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* **23**, 1801–1806. <https://doi.org/10.1093/bioinformatics/btm233> (2007).
71. Ramasamy, R. K., Ramasamy, S., Bindroo, B. B. & Naik, V. G. STRUCTURE PLOT: a program for drawing elegant STRUCTURE bar plots in user friendly interface. *SpringerPlus* **3**, 431. <https://doi.org/10.1186/2193-1801-3-431> (2014).
72. Belkhir, K., Borsa, P., Chikhi, L., Raufaste N. & Bonhomme F. GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations. Laboratoire Génome, Populations, Interactions, CNRS UMR 5000, Université de Montpellier II, Montpellier (France). (1996–2004).
73. Excoffier, L. & Lischer, H. E. L. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resour.* **10**, 546–67. <https://doi.org/10.1111/j.1755-0998.2010.02847.x> (2010).
74. Narum, S. R. Beyond Bonferroni: less conservative analyses for conservation genetics. *Conserv. Genet.* **7**, 783–787. <https://doi.org/10.1007/s10592-005-9056-y> (2006).
75. Mantel, N. The detection of disease clustering and a generalized regression approach. *Cancer Res.* **27**, 209–220 (1967).

## Acknowledgements

This work was funded by a CTM2014-53838-R Grant from the Spanish Government (Ministerio de Economía y Competitividad) and by a ED431C 2018/S7 Grant from Xunta de Galicia (Programa de Investigación Competitiva do Sistema Universitario Galego, Modalidade de grupos de referencia competitiva: Grupo de Investigación en Biología Evolutiva). I. González-Castellano was supported by a FPU scholarship from Ministerio de Educación, Cultura y Deporte (Spain), FPU15/05265. We thank Drs. Zeltia Torrecilla and Enrique González-Ortegón for kindly supplying some rockpool shrimp samples; Dr. Zeltia Torrecilla also for great ideas and discussions; Dr. Neus Marí-Mena for her kind assistance in population genetic structure analyses; and Ms. Ana M. Vázquez-Fernández for graphic support in making the figures. We also thank the anonymous reviewers for their improvements in the manuscript.

## Author contributions

I.G.C., A.M.G.T. and A.M.L. conceived the study and designed the experiments. A.M.L. collected the samples. I.G.C. performed the experiments. I.G.C. and J.G.L. analysed and interpreted the data. I.G.C. and J.G.L. prepared the figures and tables. I.G.C. wrote the manuscript. All authors reviewed drafts of the manuscript and approved its final version.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41598-020-67824-7>.

**Correspondence** and requests for materials should be addressed to I.G.-C. or A.M.-L.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020