Heterogeneous microgeographic genetic structure of the common cockle (Cerastoderma edule) in the Northeast Atlantic Ocean: biogeographic barriers and environmental factors

Vera, Manuel; Wilmes, Sophie-Berenice; Maroso, Francesco ; Hermida, Miguel; Blanco, Andrés; Casanova, Adrian ; Iglesias, David; Cao, Asunción; Culloty, S.C.; Mahony, Kate; Orvain, Francis; Bouza, Carmen; Robins, Peter; Malham, Shelagh; Lynch, Sharon; Antonio, Villalba; Martínez, Paulino

Heredity

E-pub ahead of print: 18/08/2023

Peer reviewed version

Manuscript, June 2023

Do not distribute without authors' permission

Manuel Vera [1*], Sophie B. Wilmes [2], Francesco Maroso [1], Miguel Hermida [1], Andrés Blanco [1], Adrián Casanova [1], David Iglesias [3], Asunción Cao [3], Sarah C. Culloty [4,5,6], Kate Mahony [4,5], Francis Orvain [7], Carmen Bouza [1], Peter E. Robins [2], Shelagh K. Malham[2], Sharon Lynch [4,5], Antonio Villalba [3,8,9], The Cockle's Consortium, Paulino Martínez [1*]

Heterogeneous microgeographic genetic structure of the common cockle (*Cerastoderma edule*) in the Northeast Atlantic Ocean: biogeographic barriers and environmental factors

[1] Department of Zoology, Genetics and Physics Anthropology. ACUIGEN group. Faculty of Veterinary. Campus Terra. University of Santiago de Compostela. 27002 Lugo, Spain

[2] School of Ocean Sciences, Marine Centre Wales, Bangor University, Menai Bridge, UK

[3] Centro de Investigacións Mariñas, Consellería do Mar, Xunta de Galicia. 36620 Vilanova de Arousa, Spain

[4] School of Biological, Earth and Environmental Sciences/Aquauculture and Fisheries Development Centre, University College Cork. North Mall, Cork, Ireland

[5] Environmental Research Institute, University College Cork, Cork, Ireland

[6] MaREI Centre, Environmental Research Institute, University College Cork, Cork, Ireland

[7] UNICAEN - UMR BOREA "Biologie des ORganismes et Ecosystèmes Aquatiques" MNHN, UPMC, UCBN, CNRS-7208, IRD-207, University of Caen, Caen, France

[8] Departamento de Ciencias de la Vida, Universidad de Alcalá, 28871 Alcalá de Henares, Spain

[9] Research Centre for Experimental Marine Biology and Biotechnology (PIE), University of the Basque Country (UPV/EHU), 48620 Plentzia, Basque Country, Spain

* to whom correspondence should be addressed

TLF +34 982 82 24 25

Fax +34 982 82 24 28

e-mail: manuel.vera@usc.es, paulino.martinez@usc.es

Running title: Fine-scale genetic structuring of common cockle

Word count (excluding references, tables and figures): 6,829

Abstract

Knowledge of genetic structure at the finest level is essential for conservation of genetic resources. Despite no visible barriers limiting gene flow, significant genetic structure has been shown in marine species. The common cockle (*Cerastoderma edule*) is a bivalve of great commercial and ecological value inhabiting the Northeast Atlantic Ocean. Previous population genomics studies demonstrated significant structure both across the Northeast Atlantic, but also within small geographic areas, highlighting the need to investigate fine-scale structuring. Here, we analysed two geographic areas that could represent opposite models of structure for the species: 1) the SW British Isles region, highly fragmented due to biogeographic barriers, and 2) Galicia (NW Spain), a putative homogeneous region. 9,250 SNPs genotyped by 2b-RAD on 599 individuals from 22 natural beds were used for the analysis. The entire SNP dataset mostly confirmed previous observations related to genetic diversity and differentiation, however, neutral and divergent SNP outlier datasets enabled disentangling physical barriers from abiotic environmental factors structuring both regions. While Galicia showed a homogeneous structure, the SW British Isles region was split into four reliable genetic regions related to oceanographic features and abiotic factors, such as sea surface salinity and temperature. The information gathered supports specific management policies of cockle resources in SW British and Galician regions also considering their particular socio-economic characteristics; further, these new data will be added to those recently reported in the Northeast Atlantic to define sustainable management actions across the whole distribution range of the species.

Introduction

Knowledge of genetic diversity distribution is crucial for the sustainable management and conservation of natural resources (Leary et al. 2009; Sa-Pinto et al. 2012). This distribution is affected by larval connectivity, demographic parameters and selective processes operating on species populations. Scarcity of physical barriers in marine environments is expected to promote higher connectivity among populations in comparison to terrestrial species (Waples 1998). Moreover, marine species usually show large population sizes, which along with pelagic larval stages, often lasting several weeks, facilitate population genetic homogenization across wide regions (Sa-Pinto et al. 2012; do Prado et al. 2018). Despite these general features, genetic studies on marine organisms have frequently detected genetic differentiation, even at local scales (i.e., below the geographic scale of effective dispersal of the species studied, known as chaotic genetic patchiness (CGP); see Eldon et al. 2016), which can be explained by historical and reproductive/demographic factors (e.g. high fecundity and high mortality in early life stages, sweepstakes reproductive success; see Parrondo et al. 2022), natural selection associated with environmental conditions (Vilas et al. 2015; do Prado et al. 2018; Vera et al. 2019) and oceanic features such as residual currents, bathymetry, coastline shape, upwelling, fronts, gyres and eddies (Vera et al. 2016; Coscia et al. 2020; Handal et al. 2020; Fisher et al. 2022; Vera et al. 2022).

Different types of ocean fronts have been described across the Northeast Atlantic region, encompassing tidal mixing fronts, shelf break fronts, and freshwater fronts separating estuarine freshwater and higher salinity coastal waters (Sharples and Simpson 2019). Examples of these frontal systems on the NW European Shelf include the Celtic Sea Front (NE Celtic Sea), the Irish Sea Front (NW Irish Sea), the Alderney Race (with one of the strongest current in Europe) and the Ushant Front (W English

86 Channel) (Suberg et al. 2019). These fronts may influence genetic structure acting as
87 barriers to cross-front planktonic dispersal and as conduits through along-front dispersal
88 by frontal jets, with important influences on the pelagic distribution of larvae of marine
89 species (Galarza et al. 2009). Biogeographical barriers can also limit dispersal in marine
90 environments. In the Northeast Atlantic region, Cape Finisterre, the Cornwall
91 Peninsula, the tip of Brittany, the Llyn Peninsula, and the Alderney race along Cotentin
92 Peninsula have been identified as potential barriers to the connectivity of marine
93 organisms due to their oceanographic features, including fish (Abaunza et al. 2008;
94 Larmuseau et al. 2009) and molluscs (Dupont et al. 2007; Piñeira et al. 2008; Martinez
95 et al. 2015; Handal et al. 2020; Vera et al. 2022).
96 The common cockle, *Cerastoderma edule*, is a bivalve mollusc naturally distributed
97 throughout the Northeast Atlantic coast, from Senegal, West Africa, to Norway,
98 northern Europe, where it inhabits on intertidal and shallow subtidal soft sediments
99 (Hayward and Ryland 1995). The species is commercially exploited and provides a
100 wealth of services to coastal communities mainly in Ireland, United Kingdom, France,
101 Spain and Portugal, where it is harvested (Flach and de Bruin 1994; Carss et al. 2020;
102 Jackson-Bue et al. 2022). Cockle harvest has been reduced since the 1980s (> 100,000
103 tonnes) to nowadays (~ 25,000 tonnes in 2019) due to changes in fisheries policies,
104 overfishing, variable recruitment and mass mortalities produced by pollution, climate
105 events and parasites (Villalba et al. 2014; Mahony et al. 2020; Pampin et al. 2023).
106 Furthermore, cockles are considered keystone for ecosystem due to their role as reef
107 engineers, agents of carbon sequestration and their linking between primary producers
108 and higher trophic levels (Norris et al. 1998; Carss et al. 2020). The species is dioecious
109 and can live up to 10 years displaying fast sexual maturation (reached in the first year of
110 life) and high fecundity (Honkoop and van der Meer 1998). The reproductive period

111 occurs from April to August (Malham et al. 2012), but it can be extended to September
112 in more southern European countries such as Portugal (Mahony et al. 2021), and
113 planktonic larvae can remain in the water column for 30 days facilitating widespread
114 dispersal (de Montaudouin et al. 2003; Dare et al. 2004).
115 Genetic studies throughout the natural cockle's distribution have identified three main
116 population genetic units: i) a southern group encompassing the Atlantic coast from
117 Morocco to the Bay of Biscay; ii) a central group comprising of the Celtic and Irish
118 Seas, the English Channel and the southern North Sea; and iii) a northern group
119 consisting of the northern North Sea (Beaumont et al. 1980; Hummel et al. 1994;
120 Martínez et al. 2013; 2015). These results have been recently confirmed by Vera et al.
121 (2022) through a wide genome scan (~10,000 single nucleotide polymorphisms, SNPs),
122 but additionally enabled identifying substructure within the main genetic groups using
123 outlier loci under divergent selection, mostly in accordance with residual current
124 patterns and environmental variables.
125 However, due to the limited number of markers and/or the scale of sample collection, a
126 comprehensive picture of population connectivity in the common cockle is still
127 incomplete. Information at the microgeographic level, always considering the dispersal
128 capacity of the species (Eldon et al. 2016; Vera et al. 2022), is relevant for the
129 management of fisheries (Bernatchez et al. 2017). Using a wide SNP genomic
130 screening, Coscia et al. (2020) identified three genetic clusters (global F_{ST} = 0.021) of
131 cockles in the Celtic and Irish seas and that could be associated with residual ocean
132 currents, salinity and geographical proximity using information on larval dispersal.
133 This study aimed to analyse the genetic structure of the common cockle at a
134 microgeographic scale using 2b Restriction Associated DNA sequencing (2b-RADseq).
135 Two regions were investigated: (1) the SW British Isles and the English Channel,

characterised by putative habitat fragmentation due to tidal mixing fronts and biogeographical barriers; and (2) the Northwest coast of Spain (Galicia), representing a quite homogeneous region according to previous information on other mollusc species (Diz and Presa 2009; Vera et al. 2016). The results confirmed the significant differentiation of cockles' populations at microgeographic scale, but also the power of larval dispersal to homogenize rather wide coastal areas, thus providing essential information for proper management of this valuable resource.

Material and methods

Sample area and oceanography

Two geographic areas along the Northeast Atlantic coast were investigated (Fig. 1). The first was focused on the British Isles and English Channel (hereafter called the SW British Isles region), where previous, though incomplete information, supported significant genetic sub-structuring (e.g. Coscia et al. 2020; Vera et al. 2022). The second area was Galicia (Northwest Spain), which may be genetically homogeneous according to information in other mollusc species (Diz and Presa 2009; Vera et al. 2016).

Over the cockle reproductive season (May to September; Mahony et al. 2020), the coastline of Galicia is characterised by wind-driven upwelling of cold waters resulting in sea surface temperatures (SSTs) that are several degrees colder than off-shore SSTs (Supplementary Fig. 1b). Also driven by the predominantly northerly winds in the summer months, the Portugal coastal current transports waters southwards along the coastline of Iberia (Teles-Machado et al. 2016) with residual current strengths along the Galician coastline exceeding 0.15 m/s (Supplementary Fig. 1d). The SW British Isles region is divided into distinct oceanographic regions (the English Channel, the Celtic

Deep, the Celtic Sea and the Irish Sea) by diverging current or seasonal frontal systems (Galparsoro et al. 2014). Several tidal mixing fronts separate seasonally stratified and mixed waters (Supplementary Fig. 1a): the Ushant Front (Group "Grepma", 1988), the Celtic Sea Front, and the Irish Sea Front (Simpson and Pingree, 1978). The Celtic Sea is characterised by northward flow along the western coast of Cornwall which merges into the Celtic Sea Front jet and links into the Irish Coastal Current which transports water clockwise along the south and west coast of Ireland (Supplementary Fig. 1c; Brown et al. 2003; Fernand et al. 2006). Northward currents along the Ushant Front link the Amorican Shelf with the Celtic Sea. The southern English Channel coast is dominated by northeastward flow, with the strongest currents occurring around the Cotentin Peninsula.

Sample collection

A total of 374 cockles from 14 wild natural beds were collected across the aforementioned two regions in the period 2017-2020 and stored in 100% ethanol for analyses (Table 1). Additionally, 231 cockles from eight beds previously analysed (Vera et al. 2022: identified as IDA_18, IDC_18, WDE_17, WBY_17, FBS_17, FAR_17, SNO_17 and SLO_17, where 17 and 18 in the codes represent 2017 and 2018, respectively) were included in the analysis to achieve a comprehensive picture of the areas studied, thus providing an overall total of 605 cockles. To avoid generation overlapping, all samples belonged to the 0+ year age class of their sampling year. No temporal replicates were included considering the temporal genetic stability previously reported by Vera et al. (2022).

Single Nucleotide Polymorphism (SNP) genotyping

186 Total DNA was extracted from gills using the E.Z.N.A. E-96 mollusc DNA kit
187 (OMEGA Bio-tek), following manufacturer recommendations. 2b-RAD libraries (~ 90
188 cockles per run) were constructed using the AlfI IIb restriction enzyme and sequenced
189 in an Illumina NextSeq 500 platform following Maroso et al. (2018; 2019). Bowtie
190 1.1.2 (Langmead et al. 2009) was used to align reads to the cockle's genome (Bruzos et
191 al. 2022) allowing a maximum of three mismatches and a unique valid alignment (-v 3 -
192 m 1). The reference-based mode with default parameters in the gstacks module of
193 STACKS 2.0 (Catchen et al. 2013) was used for SNP calling. For genotyping, SNPs
194 were filtered following Vera et al. (2022): *i*) SNPs genotyped in > 60% individuals; *ii*)
195 MAC (minimum allele count) ≥ 3; iii) conformance to Hardy-Weinberg expectations
196 (i.e. SNPs with significant F_{IS} values ($P < 0.05$) in at least 25% of the populations were
197 removed); and *iv*) the most polymorphic SNP within each RAD-tag was retained.
198 Individuals with less than 250,000 reads were discarded.
199
200 *Genetic diversity and population structure*
201 Estimates of genetic diversity (i.e. mean number of alleles per locus (Na), observed
202 (Ho) and expected (He) heterozygosity, proportion of polymorphic loci), departure
203 from Hardy–Weinberg equilibrium (HWE) and inbreeding coefficients (F_{IS}) were
204 estimated using GENEPOP v4.0 (Rousset 2008) and ARLEQUIN v3.5 (Excoffier and
205 Lischer 2010). Because a minimum allele frequency (MAF) filtering was not applied,
206 ARLEQUIN was also used to estimate Ho, He and F_{IS} exclusively with polymorphic
207 loci (Minimum Allele Frequency (MAF) > 0.017 according to sample size) for
208 comparison with previous studies.
209 Global and pairwise coefficients of population differentiation (F_{ST}) between cockle
210 beds were calculated with ARLEQUIN v3.5 using 10,000 permutations to test for

significance. The variational Bayesian clustering method implemented in the package fastSTRUCTURE v2.3.4 (Raj et al. 2014) was used to estimate the number of genetic population units (K) in the whole studied area and in each region testing from K = 1 to K = number of beds + 1, with an admixture ancestry model, convergence criterion of 1 x 10^{-8}, five cross-validated sets and the simple prior (flat-beta prior). The most likely number of K was estimated using the "chooseK.py" program included in the fastSTRUCTURE which gives the best K value and the K corresponding with weak population structure in the data using heuristic scores. Summarised outputs were carried out using the software POPHELPER (Francis 2017). Discriminant analyses of principal components (DAPC) were run in ADEGENET package (Jombart et al. 2010; Jombart and Ahmed 2011) for the R platform (R Development Core Team, 2014; http://www.r-project.org) with the whole dataset and for each region. Data were transformed using PCA (Principal Component Analysis) and the optimal number of principal components (PC) was calculated using the optim.a.score() command (see Miller et al. 2020). Isolation by distance (IBD) was checked by the correlation between geographical (measured as the shortest oceanic distance between two beds in Km) and genetic distance (measured as $F_{ST}/1-F_{ST}$; Rousset 1997) matrices using a Mantel test with 10,000 permutations using NTSYS v.2.1 (Rohlf 1993).

Outlier tests

The Bayesian F_{ST} –based method implemented in BAYESCAN v2.1 (Foll and Gaggiotti 2008) was used to identify outlier loci subjected to selection. BAYESCAN was run using default parameters (i.e. 20 pilot runs; prior odds value of 10; 100,000 iterations; burn-in of 50,000 iterations and a sample size of 5,000, hereafter "BY10"), but we also explored increasing prior odds value to 1000 (hereafter "BY1000"). Despite

high prior odds tend to remove false positives, they also reduce the power for detection loci under selection (Foll 2012). Loci with a False Discovery Rate (FDR, q-value) < 0.05 were considered as outliers. Moreover, the principal components-based method implemented in R package PCADAPT v4.0 (Luu et al. 2017; Prive et al. 2020) was also applied. This method renders low false-positive rates and uses individual information, not requiring *a priori* population assignment. For the analysis, the number of principal components (PC) retained was performed with the "chooseK" option. The outlier identification was carried out with an FDR < 0.05. We considered as outliers those loci identified by any of the two approaches, but additionally those shared between all approaches as the most confident ones.

Seascape analyses

Effects of spatial (latitude and longitude) and relevant abiotic factors in coastal and marine environments (sea surface temperature (SST, ºC); sea bottom temperature (SBT, ºC); sea surface salinity (SSS, psu); sea bottom salinity (SBS, psu); bottom shear stress (BSS, $N \cdot m^{-2}$); net primary productivity (NPP, $mg \cdot m^{-3} \cdot day^{-1}$); see Coscia et al. 2020 and Vera et al. 2022) shaping genetic differentiation across beds in the studied areas were assessed using a canonical redundancy analysis (RDA) implemented in the VEGAN software (Oksanen 2015) in R. This abiotic information was retrieved as monthly averages from the IBI_REANALYSIS_PHYS_005_002 ocean reanalysis model (https://resources.marine.copernicus.eu/?option=com_csw&task=results?option=com_csw&view=details&product_id=IBI_REANALYSIS_PHYS_005_002") and IBI_REANALYSIS_BIO_005_003 model (https://resources.marine.copernicus.eu/?option=com_csw&task=results?option=com_csw&view=details&product_id=IBI_REANALYSIS_BIO_005_003) for the period

2014-2018 (Supplementary Table 1), respectively. The nearest model cell classified as ocean was selected to extract the data (average distance between the sampling location and centre of the nearest model grid cell edge = 11.6 km). Then, averages for the spawning season (i.e., from April to September, see Malham et al. 2012; Mahony et al. 2020), winter (i.e., from January to March) and summer (i.e., from July to September), were calculated for each bed. Allele frequencies were calculated for each bed with ADEGENET package using the "makefreq" option. Loci with missing values were removed from the analysis. The significance of the variance associated to the different variables was tested with 1,000 random permutations. Variance inflation factor (VIF) was estimated to explore collinearity (correlation) between seascape variables in the dataset, with VIF values > 10 suggesting important collinearity problems (Marquardt, 1970). The selection model was performed using automatic stepwise model building algorithm based on permutation p-values tests. This procedure was performed with the *ordistep* function included in VEGAN. The reduced panel of explanatory variables was used to recalculate the total proportion of genetic variation in the variance partitioning. The weight of the different loci on the significant environmental vectors was calculated using VEGAN. All these analyses were performed separately for the whole, neutral and divergent outlier SNP datasets in the regions studied.

Potential correlations between allele frequencies and seascape variables were investigated with BAYENV2 (Coop et al. 2010; Gunther and Coop 2013) and results were compared with the mentioned RDA analyses. The method implemented in this software allows controlling the neutral genetic structure, because the fit improvement for a given genetic variant between a model including the environmental factor and a model including only neutral genetic structure is tested (Rellstab et al. 2015). BAYENV2 was carried out using the whole SNP datasets from SW British Isles and

Galicia, respectively. First, analyses were performed with 100,000 iterations across five independent runs to obtain the average covariance matrix for each subset. Secondly, the correlation between each SNP and the different variables was calculated using 100,000 iterations to obtain Bayes factors (BF). As in the previous step, five independent runs were used. Only SNPs with a BF > 10 and Spearman's coefficient (rho, ρ) thresholds > 1% for any variable in all runs were considered as a well-supported environment-associated SNPs. Finally, significantly correlated SNPs were compared with the outliers identified in the BAYESCAN and PCADAPT analyses.

Gene mining and functional enrichment

RAD-tags including divergent outlier SNPs were mapped in the *C. edule* genome (Bruzos et al. 2022) and their position compared with the consistent genomic windows under divergent selection previously reported by Vera et al. (2022) in the Northeast Atlantic Ocean. The very low genetic differentiation with neutral markers in the studied areas precluded the detection of consistent genomic regions under stabilizing selection. Thus, we could verify in more restricted geographical scenarios (SW British Isles and Galicia) the consistency of the genomic regions under divergent selection previously detected. Additionally, we looked for new regions under selection considering the singularity of the new sample collections of this study following a similar methodology to that proposed by Vera et al. (2022). Briefly, we defined a consistent window when ≥ 2 consecutive outliers were detected; then, we expanded the region ± 250 kb from the external outliers of the seed to define a genomic window for mining. Genes included in those genomic windows were identified using the cockle's transcriptome assembled and annotated by Pardo et al. (2022), which was used as reference to detect Gene Ontology (GO) functional enrichment of the genomic regions under selection (FDR 5%) using GOfuncR (Grote

311 2022). Furthermore, we also analysed genomic windows around the SNPs correlated with
312 environmental variables for mining; since we could not identify consecutive SNPs as with
313 outliers, we were more conservative and defined smaller windows around each SNP (±
314 100 kb).

315

316 **Results**

317 *Genetic diversity and differentiation: whole sample and SNP dataset*

318 A total of 599 cockles were analysed, since six specimens that exhibited a low number
319 of reads (< 250.000 reads) from WDE_17 (two individuals), SAN_17 (one individual),
320 SVI_17 (two individuals) and SMO_17 (one individual), were removed. After quality
321 filtering, the number of SNPs retained in the whole dataset was 9,250. This number was
322 slightly lower than the number used in the macrogeographical study carried out by Vera
323 et al. (2022) (9,309 markers), because 59 of these markers were monomorphic in the
324 studied regions. All the 9,250 markers were included in the "9,309 markers" dataset
325 and their genomic information is available at
326 https://onlinelibrary.wiley.com/doi/10.1111/eva.13340, where the SNP code from Vera
327 et al. (2022) has been maintained for comparison between studies.
328 Observed (Ho) and expected (He) heterozygosities ranged respectively from 0.070
329 (SMO_17, Spain) to 0.080 (IWC_20, IGC_20 and IKF_20, Ireland; mean ± SD = 0.075
330 ± 0.003) and from 0.076 (WDE_17, Wales) to 0.087 (SNO_17, Spain and IKF_20,
331 Ireland; mean ± SD = 0.082 ± 0.003) (Table 1). All F_{IS} values per locus and bed were
332 positive, suggesting heterozygote deficit, but low (always < 0.115) and not
333 significant;and all beds met to HW expectations ($P < 0.0022$; 0.05/22 populations), an
334 expected outcome considering the HW filtering applied to retain SNPs. The percentage
335 of polymorphic loci ranged from 25.5 % in SMO_17 (Spain) to 52.8 % in SNO_17

(Spain) (mean ± SD = 41.3 ± 7.2%). When only polymorphic loci within each bed were considered, Ho ranged from 0.137 in SLO_17 (Spain) to 0.181 in SMO_17 (Spain) (mean ± SD = 0.155 ± 0.014), showing these two beds also the lowest (0.156) and highest (0.200) He (mean ± SD = 0.171 ± 0.012). No differences in genetic diversity were found between the SW British Isles and Galician regions (Mann-Whitney U tests $P > 0.250$ for Ho, He with all loci and with polymorphic loci). Genetic diversity was in the range of previous values reported by Vera et al. (2022) for the whole Atlantic area using the same methodology.

Global F_{ST} for all beds was 0.02118 ($P < 0.001$), pairwise F_{ST} ranged from 0 (non-significant ≠ 0) for many bed pairs up to a maximum of 0.05040 ($P < 0.001$) between IDC_18 and SVI_17 (Supplementary Table 2). Most pairwise comparisons were significant excluding those from Galicia. Average pairwise F_{ST} between the SW British Isles and Galicia was 0.03400 ($P < 0.001$), while 0.01374 ($P < 0.001$) within the SW British Isles and -0.00529 ($P = 1.000$) within Galicia. The two beds from the Cotentin Peninsula (FBV_19 and FGO_19, SW British Isles region), separated by 190 kilometres, showed significant genetic differentiation ($F_{ST} = 0.01207$, $P < 0.001$). The most likely K values inferred by fastSTRUCTURE were 1 and 3. When K = 3 was plotted, two main groups were identified differentiating the SW British Isles (IGC_20, IKF_20, IWC_20, IDA_18, IDC_18, WDE_17, WBY_17, ECE_20, FBS_17, FBV_19 and FGO_19) from Galicia (plus Arcachon) (FAR_17, SBA_17, SMI_17, SAN_17, SNO_17, SLO_17, SSA_17, SVI_17, SCA_17, SMO_17, SBI_18) (Fig. 2A). FGO_19 (France, Cotentin Peninsula) showed a high component of the southern group, also detectable in all samples from the English Channel (ECE_20, FBS_17 and FBV_18), suggesting some introgression between the two groups. The DAPC representation on the SW British Isles also suggested differentiation of the English Channel samples from

361 the northernmost populations across the second component, while the first one,
362 indicated a remarkable divergence of the IKF_20 sample from the remaining ones (Fig.
363 3A). The DAPC from Galicia showed most of the samples grouped excluding SNO_17,
364 in the middle of the distribution, below Cape Finisterre, and SBI_18, the southernmost
365 one (Fig. 4A).
366
367 *Genetic structure within regions: demographic and selective factors*
368 To understand the factors underlying genetic differentiation within the SW British Isles
369 and Galician regions, we first identified those loci under selection using three different
370 statistical approaches. BY10, BY1000 and PCADAPT detected 159, 47 and 84 outliers
371 in the SW British Isles, respectively, all of them under divergent selection and
372 representing a total of 186 different outliers (Supplementary Table 3). Thirty-five
373 markers were shared between the three methods. The number of outliers in Galicia was
374 much lower (BY10 = 15, BY1000 = 2, PCADAPT = 39), two of them shared between
375 the three methods and representing a total of 51 outliers, all of them putatively under
376 divergent selection. Among the whole outlier dataset, 15 were shared between SW
377 British Isles and Galicia. Then, by discounting the total number of outliers to the whole
378 dataset in each region, a total of 9,064 neutral markers were identified in the SW British
379 Isles and 9,199 in Galicia, representing the neutral datasets for each region.
380 Small but significant genetic differentiation was detected among the SW British Isles
381 beds using neutral markers (F_{ST} = 0.00778, $P < 0.001$), suggesting limitations to larval
382 dispersion in this area by biogeographical barriers. As expected, the 186 total outliers
383 rendered a much higher global F_{ST} (0.10959, $P < 0.001$), being more accentuated when
384 using the shared set of outliers between methods (F_{ST} = 0.17411, $P < 0.001$), which
385 suggests selective factors increasing structuring. Pairwise F_{ST} ranged from -0.03095

(IKF_20 – WDE_17) to 0.02366 (IGC_20 – FGO_19 pair) for neutral markers; from 0.00061 (IDA_18 – IDC_18 pair) to 0.17142 (IDA_18 – FGO_19 pair) for the 186 total outliers; and from -0.00345 (IDA_18 – IDC_18 pair) to 0.27313 (IKF_20 – FBS_17 pair) for the 35 shared outliers (Supplementary Table 4). IBD was significant with the shared and total outlier datasets ($r = 0.63169$ and 0.55136, respectively; $P < 0.001$), but not with the neutral dataset ($r = 0.08578$, $P = 0.330$). These results suggest that correlations could be a by-product of the unequal spatial distribution of the environmental factors responsible of selective forces shaping the cockle's genome, since IBD patterns should be reflected by the balance between drift and migration on neutral markers. The fastSTRUCTURE analyses identified $K = 1$, $K = 2$ and $K = 3$ as the most likely values for the neutral, 35 shared outlier and 186 total outlier datasets, respectively (Fig. 2B), which consistently differentiated the Celtic Sea and the North-west Irish cluster (IGC_20, IKF_20 and IWC_20), not studied to date, and the English Channel cluster (ECE_20, FBS_17, FBV_19, FGO_19). In contrast, the Irish Sea appeared as a rather differentiated group with the 186 outliers, which was split into two clusters, the Irish side (IDA_18 and IDC_18) most closely associated with the Celtic and Atlantic Ocean cluster, and the Welsh side (WDE_17 and WBY_17), most closely linked with the English Channel cluster, when using the 35 outlier loci. The differentiation of the Irish Sea from the other samples, and the contrast between the Welsh and Irish (east and west, respectively) samples of the Irish Sea, was shown when exploring a scenario with a larger K value, with both datasets displaying a very similar structure with $K = 4$ (Supplementary Figures 2 and 3). The DAPC analysis with neutral markers showed a very similar picture to that described with the whole SNP dataset (Fig. 3B), however, the 186 and 35 outlier datasets displayed a very distinct picture, both separating the English Channel (ECE_20, FGO_19, FBV_19 and FBS_17) from

411 the Welsh populations, but also from the Irish populations, which were further divided
412 into two groups, the westernmost Northeast Atlantic Ocean group (IWC_20, ICG_20
413 and IKF_20) and the Irish/Celtic Seas group (IDC_18 and IDA_18) (Figs. 3C and 3D).
414 In contrast to the SW British Isles, no population differentiation was found in Galicia
415 with the neutral dataset (F_{ST} = 0.00552, P = 1.000), also supported by the
416 fastSTRUCTURE (K =1) and DAPC, as previously outlined with whole dataset (Figs.
417 4A and B). However, low but significant differentiation was detected with the 51
418 outliers (F_{ST} = 0.00870, P < 0.001), the pairwise F_{ST} supporting a significant
419 differentiation of the two northernmost samples (SMI_17 and especially SBA_17;
420 Supplementary Table 5) from the rest. This differentiation was not disclosed with
421 fastSTRUCTURE (K = 1; see Supplementary Fig. 4) and only suggested with DAPC
422 (Fig. 4C).
423
424 *Seascape analysis*
425 RDA analyses in SW British Isles region suggested longitude as the main driver for the
426 observed differentiation with all datasets and seasons (Table 2). Latitude was also
427 supported as driver for many models, especially for those related to the 186 total
428 outliers. Sea bottom salinity (SBS) was suggested for all seasons with the 186 outlier
429 dataset, while bottom shear stress (BSS) was for reproductive and summer seasons
430 using the whole and neutral datasets (Table 2). When longitude and latitude were
431 removed, sea surface temperature (SST) was suggested for all the datasets in the
432 summer season, and in the reproductive and winter seasons only with the whole and
433 186 outlier datasets, respectively. Sea bottom temperature (SBT) was suggested for the
434 reproductive and summer season with the 186 outlier dataset. SBS and sea surface
435 salinity (SSS) were suggested with the 186 outlier dataset for the summer and winter

436 seasons, respectively. Net primary production (NPP) was suggested for all datasets in
437 the winter season and for the 186 outlier dataset for the reproductive season. Finally,
438 BSS was suggested in all seasons for the neutral dataset and in the reproductive and
439 summer seasons for the complete dataset. In Galician region, no associations were
440 found, except for latitude in all periods analysed using the 51 outliers, and for BSS
441 during winter when latitude and longitude were removed (Table 2). However, VIF
442 values were usually high (> 10), suggesting that results should be taken with caution
443 due to the high collinearity among the variables in many cases.
444 While no correlations were identified in Galicia with BAYENV2, a total of 54 markers
445 were correlated with different environmental variables in the SW British Isles
446 (Supplementary Table 6). Thirty of these markers (55.6%) were previously identified as
447 outliers by the different methodologies applied. Markers were mainly correlated with
448 latitude, longitude, temperature and salinity. The main variable correlated with genetic
449 markers in the reproductive season and summer scenarios was SBT, while SSS and
450 NPP were in the winter scenario.
451
452 *Gene mining around outliers and environmental correlated markers*
453 Genetic markers associated with divergent selection or correlated with environmental
454 variables were mapped in the common cockle genome to look for functional
455 interpretation (Supplementary Tables 3 and 6). Outliers identified in the SW British
456 Isles area were scattered across all chromosomes, between one in C18 and 22 in C3,
457 while five chromosomes (C8, C11, C13, C14, C17) did not bear any outlier in Galicia,
458 the maximum being detected in C1 (11 outliers) (Table 3). The 51 outliers detected in
459 Galicia only identified a single consistent genomic region (window) under selection
460 according to our criteria and other five outliers were distributed across four confident

genomic windows previously reported by Vera et al. (2022) (Supplementary Tables 7 and 8). However, among the 186 outliers detected in the SW British Isles, 14 defined five new consistent genomic windows under divergent selection and other 45 mapped on genomic windows previously reported by Vera et al. (2022) (Supplementary Tables 7 and 8). Most outliers detected in Galicia were specific to this region, while an important number of outliers from the SW British Isles were shared with the Northern region previously analysed by Vera et al. (2022) (Supplementary Fig. 5). Still, a notable proportion of outliers in the North were specific of each study (North-Vera et al. (2022): 137 vs SW British Isles: 101) suggesting specific evolutionary factors related to each scenario. Among the genes annotated in the five new windows, several related to oxidative stress, hypoxia and immunity were identified in a 200 kb region in C2 and in a 340 kb region in C3 (Supplementary Table 9) (Gerdol and Venier 2015; Grandi et al. 2016; Sokolov et al. 2019). Also, in a 480 kb region in C5, some genes involved in signalling and detoxification (Wang et al. 2018; Kron 2022; Thoma et al. 2022) were identified. Finally, a gene associated with ocean acidification (Lim et al. 2021) was identified in C19. Despite the low number of genes handled, a significantly enriched GO Molecular Function was detected (protein serine/threonine phosphatase activity; GO:0004722) taking as background the common cockle transcriptome reported by Pardo et al. (2022).

Markers correlated with environmental variables were scattered across most chromosomes, excluding C7, C14 and C18, and the higher number (seven markers) were detected in two big chromosomes (C2 and C4) (Table 3). An important number of correlated markers were also identified as outliers for divergent selection (55.6 %), some of them associated with consistent genomic windows (Supplementary Table 8). Of note, the three markers detected in one of the most consistent genomic windows in

C4. We also mined the cockle genome around the correlated marker dataset (Supplementary Table 10) and detected several genes related to nervous system development and physiology. These genes were mostly clustered at C1 around 142462_31 (correlated with SBT) and C2 around 210318_7 (correlated with SST), respectively. Furthermore, some of these genes were previously associated with temperature stress and oxygen depletion stress or differentially expressed under specific experimental conditions in other mollusc species (Woo et al. 2011; Chen et al. 2022). Another important group of genes scattered around different markers in the cockle genome were related to immunity and defence and had been previously reported in other mollusc species in response to viruses and bacteria (Barbosa et al. 2022; Saco et al. 2023) (Supplementary Table 9).

Discussion

Assessment of the distribution of genetic variability across populations, incorporating historical processes and local adaptation framed within the dispersal range of the focal organism (Richardson et al. 2014), is essential to develop management actions to preserve exploited species (Bernatchez et al. 2017). In the present study, two different patterns of genetic structure at microgeographic scale were identified in two regions within the natural distribution of *C. edule*, highlighting the need to perform analyses at multiple spatial scales (Hoffman et al. 2012), to provide information supporting the management of this valuable resource.

Heterogeneous pattern of microgeographic structure in the common cockle

The two geographic areas studied, the SW British Isles region and Galicia, were selected by their different habitat fragmentation patterns. Both areas were slightly differentiated (F_{ST} = 0.03400), in accordance with their location in the major northern and southern regions of the species' range separated around French Brittany (Vera et al.

2022), but did not show differences in genetic diversity, unlike Vera et al. (2022), who reported a slight, but significant higher diversity in the southern region.

The extensive analysis performed in Galicia (10 natural beds) suggested the presence of a single panmictic unit in this area, as previously reported for other molluscs, with similar pelagic larval periods (*Donax trunculus*, Nantón et al. 2017; *Ensis siliqua*, Arias-Pérez et al. 2012; *Mitilus galloprovincialis*, Diz and Presa 2009; *Ostrea. edulis*, Vera et al. 2016; *Polititapes rhomboides*, Chacón et al. 2021), and for other marine species (*Hippocampus guttulatus*; Lopez et al. 2015; *Pollicipes pollicipes*, Parrondo et al. 2022). Our data does not support Cape Finisterre as a biogeographical barrier for the species as previously suggested (Lopez-Jamar et al. 1992; Piñeira et al. 2008; Martinez et al. 2013; Cruz et al. 2020), since no differentiation was detected between beds at both sides of the Cape with the whole and neutral datasets. However, when using outlier loci, the two northernmost Galician beds showed significant differentiation with the remaining ones, especially the bed closest to the Cantabrian Sea (SBA_17) (average F_{ST} = 0.03145), which could be related to the higher temperature regime in the Cantabrian Sea (Marquina et al. 2015), but a more detailed study in the Cantabrian Sea would be necessary to confirm this observation. Oceanographic dynamics in the Galician coast indicate that the cold-upwelled water usually penetrates estuaries on the west, while it only occurs during very intense events on the north (Alvarez et al. 2010). Thus, water temperature decreases from north to west, with an SST average value of 19.5 ºC in the Cantabrian coast compared with the 18.5 ºC measured in the west coast for the 1985-2005 period (Gomez-Gesteira et al. 2008). Larval dispersal modelling carried out in Vera et al. (2022) (see their Fig. 7) confirmed that cockle beds are well connected with each other by larval transport in Galicia, but the connection between the Rias and the sites to the northeast of Cape Finisterre was weaker, though present. Furthermore,

whilst the beds along the northwest coast of the Iberian Peninsula are affected by very similar oceanographic conditions, during the late spring and late summer, temperatures at the most north-easterly site can differ markedly from those at the other beds due to its location at the edge of the upwelling system and at the inception point of the Portugal Coastal Current (STT two degrees higher in the northern beds (mean = 18.51 ºC) than in the southern ones (mean = 16.47 ºC) during the summer; see Supplementary Table 1). Despite the genotype-environment associations methods did not identify sea temperature as driver, latitude, which is highly correlated with temperature, was suggested by the RDA analysis as potential driver in the region.

Previous data from the SW British Isles suggested significant structure in *C. edule* related both to current dynamics as well as to abiotic factors, such as salinity and temperature (Coscia et al. 2020; Vera et al. 2022), as reported in other shellfish species such as the horse mussel *Modiolus modiolus* (Gormley et al. 2015) and the great scallop *Pecten maximus* (Vendrami et al. 2019; Hold et al. 2021). However, some regions in this area are still poorly sampled in the common cockle (English Channel) or without information (West Irish coast, North-east Atlantic). Outlier markers showed a moderate pairwise genetic differentiation between beds (F_{ST} = 0.10959 and 0.17411 with the 186 and 35 outlier datasets, respectively), higher than that observed with neutral markers (F_{ST} = 0.00778), as expected, suggesting selective factors shaping specific genomic regions in a small geographic area. An important proportion of divergent outliers (68 markers) were shared with those reported by Vera et al. (2022) for the northern group (210 outliers), which gives robustness to our observations; however, data also suggests specific selective factors shaping the cockle's genome associated with the new sampling in the SW British Isles (117 new outlier loci; 31 within consistent genomic windows). In fact, five new confident genomic windows were identified including

561 relevant genes related to oxidative stress and immunity that would deserve further
562 studies as candidates to explain the association observed with environmental factors.
563 Despite biotic factors, such as pathogens, could not be contemplated in our study, their
564 diversity and distribution (influenced by abiotic factors) are important drivers shaping
565 the genome and distribution of species (Theodosopoulos et al. 2019) and specifically in
566 cockles (Vera et al. 2022; Pampín et al. 2023). Furthermore, we also deepened into the
567 correlation of specific SNPs with environmental factors and could identify, by mining
568 in the cockle genome several genes related to nervous transmission and immunity,
569 arranged in clusters or scattered in different chromosomes, that had been previously
570 reported in other mollusc associated with temperature or oxidative stress (Woo et al.
571 2011; Barbosa et al. 2022; Chen et al. 2022).
572 The population structure observed in the SW British Isles region may be in part
573 explained by the residual ocean currents and ocean fronts that characterise this area, but
574 also by selective factors such as salinity gradients, variable bottom shear stress (due to
575 large tidal variability) and sea temperature gradients (driven by ocean currents,
576 stratification and mixing, and latitudinal gradients); however, spatial seascape results
577 should be taken with caution due to the collinearity detected among variables. Both
578 outlier datasets could identify four genetic clusters following two main west-east and
579 north-south axes, which could explain the correlation observed between genetic and
580 geographic distances for outlier loci, but also the identification of longitude and latitude
581 as two main drivers in the seascape analysis. According to the outlier information, the
582 new sampled beds from Western Ireland (IGC_20, IKF_20 and IWC_20) (Northeast
583 Atlantic) would constitute a new cluster. These sites are connected by the Irish coastal
584 current (Brown et al. 2003; Fernand et al. 2006) and larval dispersal modelling (see Fig.
585 7 Vera et al. 2022) showed that the beds along the southwest coast of Ireland are well

586 interconnected. The Irish Sea can be split into two different clusters associated with the
587 Irish and Welsh sides, as previously suggested by Coscia et al. (2020). Sites along the
588 southeast coast of the Irish Sea are generally connected by northward currents and sites
589 along the north coast of Wales by eastward currents. In contrast, the two sites on the
590 west coast of the Irish Sea (IDC_18 and IDA_18), appear genetically separated from
591 the remainder of the Irish Sea; This may be driven by the Irish Sea Front acting as a
592 barrier which also drives warmer temperatures in the northwest Irish Sea than in the
593 well-mixed northeast Irish Sea. Finally, the English Channel forms a fourth cluster
594 including the ECE_20 bed from Cornwall with the southern beds limited by the Ushant
595 front. Interestingly, the Cotentin Peninsula, previously identified as a physical barrier to
596 dispersal in other molluscs, such as the slipper limpet *Crepidula fomicata* (Dupont et al.
597 2007) and *P. maximus* (Nicolle et al. 2016; Handal et al. 2020), showed a significant
598 differentiation between samples on its west and east sides (FGO_19 and FBV_19) with
599 neutral markers (F_{ST} = 0.01045, P < 0.001) and higher with outlier loci (F_{ST} 35 outliers
600 = 0.06654, P < 0.001; F_{ST} 186 outliers = 0.05876, P < 0.001), suggesting additional
601 selective factors differentiating both sides. Oceanic distance between the two Cotentin
602 beds (~ 190 km) is shorter than the longest distance between Galician beds (~ 300 km),
603 where no genetic differentiation was detected with neutral markers. Of note, FGO_19
604 showed an important genomic component of the South group, suggesting introgression
605 from the south especially in the west coast of the Cotentin Peninsula.
606
607 *Management implications*
608 The present study represents a refined analysis of the population structure of *C. edule* in
609 two geographic areas of small-medium size representing differentiated models that could
610 aid to obtain a more comprehensive picture for improving the management and

conservation of this valuable commercial and ecological resource. Galician beds were suggested to constitute a panmictic population and this region could be managed as a single genetic unit. The fisheries in this region are exclusively commercial and their exploitation management can be through territorial concessions leased by shellfisher guilds or directly by Galician regional government (i.e. free access shellfish areas). This genetic information should be included in the Galician legislation, thus allowing translocations from high production areas (Ría de Noia) to depleted ones by different factors, such as the parasite *M. cochillia* (Ría de Arousa; Villalba et al. 2014). However, caution should be taken considering biotic factors not evaluated in our study, such as emergent pathologies (e.g. marteiliosis), which will require specific recommendations within the general framework depicted in our study. A sharp fragmentation was displayed by the SW British Isles region, especially with divergent outliers, mostly representing adaptive management units (AMU, Bernatchez et al. 2017). Thus, Western (Northeast Atlantic) Irish beds would represent a differentiated group from those previously described, while subtle genetic sub-structuring was identified along the English Channel, with a significant effect at the Cotentin Peninsula representing as a biogeographic barrier. Furthermore, the Irish Sea, a narrow water body mass between Wales and Ireland, appears to represent differentiated units at both sides of the Irish Sea according to our information. All these population units should be individually managed, avoiding translocations between them. Finally, our results could help to improve cockles' production by founding appropriate broodstock to enhance depleted populations and by tracing samples to check undesirable transferences among regions.

Acknowledgements

Authors wish to thank L. Insua, S. Sánchez-Darriba and S. Gómez from the ACUIGEN group (USC) for their technical support. Supercomputing Center of Galicia

(http://www.cesga.es) provided computing facilities for genotyping. A. Casanova (ACC) was funded by a Xunta de Galicia-Campus Terra postdoctoral fellow.

Author Contributions

MV, AV and PM designed and supervised the study. DI, AC, KM, FO, SKM, SL performed field collections. PM, SCC, SL, PER, SKM and FO provided funding. FM, MH and AB analysed bioinformatically genomic sequences and created genotyping files. SBW and PER provided information about oceanography, environmental variables and developed geographic maps included in the figures. MV, CB, ACC, AB and PM performed the genetic analyses. MV and PM wrote the manuscript with contributions from all authors. All of them read the manuscript and gave their approval.

Funding

This study has been supported by the COCKLES project (grant number: EAPA_458/2016) of the INTERREG EUROPEAN PROGRAMME and the NERC-SHEAR project (NE/W001217/1).

Conflict of interest

The authors declare no conflict of interest.

Data archiving

Data for this study are available at Dryad Digital Repository (http://*to be completed after manuscript is accepted for publication)* and Supplementary material.

References

660 Abaunza P, Murta AG, Campbell N, Cimmaruta R, Comesana AS, Dahle G, Santamaria
661 MTG, Gordo LS, Iversen SA, MacKenzie K, Magoulas A, Mattiucci S, Molloy
662 J, Nascetti G, Pinto AL, Quinta R, Rarnos P, Sanjuan A, Santos AT, Stransky C,
663 Zimmermann C (2008) Stock identity of horse mackerel (*Trachurus trachurus*)
664 in the Northeast Atlantic and Mediterranean Sea: Integrating the results from
665 different stock identification approaches. Fish. Res. 89:196-209

666 Alvarez I, Gomez-Gesteira M, DeCastro M, Gomez-Gesteira JL, Dias JM (2010)
667 Summer upwelling frequency along the western Cantabrian coast from 1967 to
668 2007. J. Mar. Syst. 79:218-226

669 Alvarez I, Gomez-Gesteira M, deCastro M, Lorenzo MN, Crespo AJC, Dias JM (2011)
670 Comparative analysis of upwelling influence between the western and northern
671 coast of the Iberian Peninsula. Cont. Shelf Res. 31:388-399

672 Arias-Perez A, Fernandez-Tajes J, Gaspar MB, Mendez J (2012) Isolation of
673 microsatellite markers and analysis of genetic diversity among East Atlantic
674 populations of the sword razor shell *Ensis siliqua*: a tool for population
675 management. Biochem. Genet. 50: 397-415

676 Barbosa M, Schwaner C, Pales Espinosa E, Allam B (2022) A Transcriptomic analysis
677 of phenotypic plasticity in *Crassostrea virginica* larvae under experimental
678 acidification. Genes 13: 1529

679 Beaumont AR, Day TR, Gade G (1980) Genetic variation at the octopine
680 dehydrogenase locus in the adductor muscle of *Cerastoderma edule* (L) and 6
681 other bivalve species. Mar.Biol. Letters 1: 137-148

682 Bernatchez L, Wellenreuther M, Araneda C, Ashton DT, Barth JMI, Beacham TD,
683 Maes GE, Martinsohn JT, Miller KM, Naish KA, Ovenden JR, Primmer CR,

Suk HY, Therkildsen NO, Withler RE (2017) Harnessing the Power of Genomics to Secure the Future of Seafood. Trends Ecol. Evol. 32:665-680

Brown J, Carrillo L, Fernand L, Horsburgh KJ, Hill AE, Young EF, Medler KJ (2003) Observations of the physical structure and seasonal jet-like circulation of the Celtic Sea and St. George's Channel of the Irish Sea. Cont. Shelf Res. 23:533–561

Bruzos AL, Santamarina M, Garcia-Souto D, Diaz S, Rocha S, Zamora J, Lee Y, Viña-Feas A, Quail MA, Otero I, Pequeño-Valtierra A, Temes J, Rodriguez-Castro J, Villanueva A, Costas D, Rodriguez R, Prieto T, Tomas L, Alvariño P, Alonso J, Cao A, Iglesias D, Carballal MJ, Amaral AM, Balseiro P, Calado R, El Khalfi B, Izagirre U, de Montaudouin X, Pade NG, Probert I, Ricardo F, Ruiz P, Skazina M, Smolarz K, Pasantes JJ, Villalba A, Ning Z, Ju YS, Posada D, Demeulemeester J, Baez-Ortega A, Tubio JMC (2022) The evolution of two transmissible leukaemias colonizing the coasts of Europe. Biorxiv doi10.1101/2022.08.06.503021

Carss DN, Brito AC, Chainho P, Ciutat A, de Montaudouin X, Fernandez Otero RM, Incera Filgueira M, Garbutt A, Goedknegt MA, Lynch SA, Mahony KE, Maire O, Malham SK, Orvain F, Olivier AvdS, Jones L (2020) Ecosystem services provided by a non-cultured shellfish species: The common cockle *Cerastoderma edule*. Mar. Environ. Res. 158:104931

Catchen J, Hohenlohe PA, Bassham S, Amores A, Cresko WA (2013) Stacks: an analysis tool set for population genomics. Mol. Ecol. 22:3124-3140

Chacon GM, Arias-Perez A, Freire R, Martinez L, Ojea J, Insua A (2021) Genetic characterization of wild, broodstock and seed samples of *Polititapes rhomboides*

(Bivalvia: Veneridae): Implications for hatchery seed production. Aquacult. Rep. 20:100658

Chen J, Leng T, Jiang YM, Chen XB, Liu ZM (2022) RNA-seq analysis of the differential response to low-temperature stress in two morphs of mud crabs (*Scylla paramamosain*). Comp. Biochem. Physiol. Part D Genomics Proteomics 43:101010

Coop G, Witonsky D, Di Rienzo A, Pritchard JK (2010) Using environmental correlations to identify loci underlying local adaptation. Genetics 185:1411-1423

Coscia I, Wilmes SB, Ironside JE, Goward-Brown A, O'Dea E, Malham SK, McDevitt AD, Robins PE (2020) Fine-scale seascape genomics of an exploited marine species, the common cockle *Cerastoderma edule*, using a multimodelling approach. Evol. Appl. 13:1854-1867

Cruz A, da Costa F, Fernandez-Perez J, Nanton A, Fernandez-Boo S, Insua A, Mendez J (2020) Genetic variability in *Ruditapes decussatus* clam combined with *Perkinsus* infection level to support founder population selection for a breeding program. PeerJ 8:e9728

Dare PJ, Bell MC, Walker P, Bannister RCA (2004) Historical and current status of cockle and mussel stocks in The Wash. CEFAS.

de Montaudouin X, Bachelet G, Sauriau PG (2003) Secondary settlement of cockles *Cerastoderma edule* as a function of current velocity and substratum: a flume study with benthic juveniles. Hydrobiologia 503:103-116

Diz AP, Presa P (2009) The genetic diversity pattern of *Mytilus galloprovincialis* in Galician Rias (NW Iberian estuaries). Aquaculture 287:278-285

731 do Prado FD, Vera M, Hermida M, Bouza C, Pardo BG, Vilas R, Blanco A, Fernandez
732 C, Maroso F, Maes GE, Turan C, Volckaert FAM, Taggart JB, Carr A, Ogden
733 R, Nielsen EE, Martinez P, Aquatrace C (2018) Parallel evolution and
734 adaptation to environmental factors in a marine flatfish: Implications for
735 fisheries and aquaculture management of the turbot (*Scophthalmus maximus*).
736 Evol. Appl. 11:1322-1341

737 Dupont L, Ellien C, Viard F (2007) Limits to gene flow in the slipper limpet *Crepidula*
738 *fornicata* as revealed by microsatellite data and a larval dispersal model. Mar.
739 Ecol. Prog. Ser. 349:125-138.

740 Eldon B, Riquet F, Yearsley J, Jollivet D, Broquet T (2016) Current hypotheses to
741 explain genetic chaos under the sea. Curr. Zool. 62:551-566

742 Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to
743 perform population genetics analyses under Linux and Windows. Mol. Ecol.
744 Resour. 10:564-567

745 Fernand L, Nolan GD, Raine R, Chambers CE, Dye SR, White M, Brown J (2006) The
746 Irish coastal current: A seasonal jet-like circulation. Cont. Shelf Res. 26:1775–
747 1793

748 Fisher MC, Helser TE, Kang S, Gwak W, Canino MF, Hauser L (2022) Genetic
749 structure and dispersal in peripheral populations of a marine fish (Pacific cod,
750 *Gadus macrocephalus*) and their importance for adaptation to climate change.
751 Ecol. Evol. 12:e8474

752 Flach EC, de Bruin W (1994) Does the activity of cockles, *Cerastoderma edule* (L.) and
753 lugworms, *Arenicola marina* L., make corophium-volutator pallas more

754 vulnerable to epibenthic predators: a case of interaction modification. J. Exp.
755 Mar. Biol. Ecol. 182:265-285

756 Foll M, Gaggiotti O (2008) A Genome-Scan Method to Identify Selected Loci
757 Appropriate for Both Dominant and Codominant Markers: A Bayesian
758 Perspective. Genetics 180:977-993

759 Foll M (2012) BayeScan v2.1 User Manual. Available at:
760 http://cmpg.unibe.ch/software/BayeScan/files/BayeScan2.1_manual.pdf

761 Francis RM (2017) POPHELPER: an R package and web app to analyse and visualize
762 population structure. Mol. Ecol. Resour. 17:27-32

763 Galarza JA, Carreras-Carbonell J, Macpherson E, Pascual M, Roques S, Turner GF,
764 Rico C (2009) The influence of oceanographic fronts and early-life-history traits
765 on connectivity among littoral fish species. Proc. Natl. Acad. Sci. U.S.A.
766 106:1473-1478

767 Galparsoro I, Borja A, Uyarra MC (2014) Mapping ecosystem services provided by
768 benthic habitats in the European North Atlantic Ocean. Front. Mar. Sci. 1:23

769 Gerdol M, Venier P (2015) An updated molecular basis for mussel immunity. Fish
770 Shellfish Immunol. 46:17-38

771 Gomez-Gesteira M, de Castro M, Alvaez I, Gomez-Gesteira JL (2008) Coastal sea
772 surface temperature warming trend along the continental part of the Atlantic Arc
773 (1985-2005). J. Geophys. Res. Oceans 113:C04010

774 Gormley K, Mackenzie C, Robins P, Coscia I, Cassidy A, James J, Hull A, Piertney S,
775 Sanderson W., Porter J (2015) Connectivity and Dispersal Patterns of Protected
776 Biogenic Reefs: Implications for the Conservation of *Modiolus modiolus* (L.) in
777 the Irish Sea. PLOS One 10:e0143337

778 Grandi A, Santi A, Campagnoli S, Parri M, De Camilli E, Song C, Jin B, Lacombe A,
779 Castori-Eppenberger S, Sarmientos P, Grandi G, Viale G, Terracciano L,
780 Chiarugi P, Pileri P, Grifantini R (2016) ERMP1, a novel potential oncogene
781 involved in UPR and oxidative stress defense, is highly expressed in human
782 cancer. Oncotarget 7:63596-63610

783 Grote S (2022) GOfuncR: Gene ontology enrichment using FUNC. R package version
784 1.18.0

785 Group "Grepma" (1988) A physical, chemical and biological characterization of the
786 Ushant tidal front. Int. Rev. ges. Hydrobiol. Hydrogr. 73:511–536

787 Gunther T, Coop G (2013) Robust Identification of Local Adaptation from Allele
788 Frequencies. Genetics 195:205-220

789 Handal W, Szostek C, Hold N, Andrello M, Thiebaut E, Harney E, Lefebvre G, Borcier
790 E, Jolivet A, Nicolle A, Boye A, Foucher E, Boudry P, Charrier G (2020) New
791 insights on the population genetic structure of the great scallop (*Pecten*
792 *maximus*) in the English Channel, coupling microsatellite data and demogenetic
793 simulations. Aquat. Conserv. 30:1841-1853

794 Hayward PJ, Ryland JS (1995) Handbook of the marine fauna of north-west Europe.
795 Oxford University Press.

796 Hoffman JI, Clarke A, Clark MS, Fretwell P, Peck LS (2012) Unexpected fine-scale
797 population structure in a broadcast-spawning Antarctic marine mollusc. PLOS
798 One 7:e32415

799 Hold N, Robins P, Szostek CL, Lambert G, Lincoln H, Le Vay L et al (2021) Using
800 biophysical modelling and population genetics for conservation and

management of an exploited species, *Pecten maximus* L. Fish. Oceanogr. 30:740-756.

Honkoop PJC, van der Meer J (1998) Experimentally induced effects of water temperature and immersion time on reproductive output of bivalves in the Wadden Sea. Journal of Experimental Mar. Biol. Ecol. 220:227-246

Hummel H, Wolowicz M, Bogaards RH (1994) Genetic variability and relationships for populations of *Cerastoderma edule* and of the *C. glaucum* complex. Netherlands J. Sea Res. 33:81-89

Jackson-Bue M, Brito AC, Cabral S, Carss DN, Carvalho F, Chainho P, Ciutat A, Sanchez EC, de Montaudouin X, Otero RMF, Filgueira MI, Fitch A, Garbutt A, Goedknegt MA, Lynch SA, Mahony KE, Maire O, Malham SK, Orvain F, Rocroy M, Olivier AvdS, Jones L (2022) Inter-country differences in the cultural ecosystem services provided by cockles. People Nat. 4:71-87

Jombart T, Ahmed I (2011) adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. Bioinformatics 27:3070-3071

Jombart T, Devillard S, Balloux F (2010) Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. BMC Genetics 11:94

Kron NS (2022) In search of the *Aplysia* immunome: an in silico study. BMC Genomics 23:543

Langmead B, Trapnell C, Pop M, Salzberg SL (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 10:R25

Larmuseau MHD, Van Houdt JKJ, Guelinckx J, Hellemans B, Volckaert FAM (2009) Distributional and demographic consequences of Pleistocene climate

825 fluctuations for a marine demersal fish in the north-eastern Atlantic. J. Biogeogr.
826 36:1138-1151

827 Leary D, Vierros M, Hamon G, Arico S, Monagle C (2009) Marine genetic resources: A
828 review of scientific and commercial interest. Mar. Policy 33:183-194

829 Lim Y-K, Cheung K, Dang X, Roberts SB, Wang X, Thiyagarajan V (2021) DNA
830 methylation changes in response to ocean acidification at the time of larval
831 metamorphosis in the edible oyster, *Crassostrea hongkongensis*. Mar. Environ.
832 Res. 163:105214

833 Lopez A, Vera M, Planas M, Bouza C (2015) Conservation genetics of threatened
834 *Hippocampus guttulatus* in vulnerable habitats in NW Spain: Temporal and
835 spatial stability of wild populations with flexible polygamous mating system in
836 captivity. PLOS One 10:0117538

837 Lopez-jamar E, Cal RM, Gonzalez G, Hanson RB, Rey J, Santiago G, Tenore KR
838 (1992) Upwelling and outwelling effects on the benthic regime of the
839 continental-shelf off Galicia, NW Spain. J. Mar. Res. 50:465-488

840 Luu K, Bazin E, Blum MGB (2017) Pcadapt: an R package to perform genome scans
841 for selection based on principal component analysis. Mol. Ecol. Resour. 17:67-
842 77

843 Mahony KE, Lynch SA, Egerton S, Cabral S, de Montaudouin X, Fitch A, Magalhaes
844 L, Rocroy M, Culloty SC (2020) Mobilisation of data to stakeholder
845 communities. Bridging the research-practice gap using a commercial shellfish
846 species model. PLOS One 15:0238446

847 Mahony KE, Lynch SA, Egerton S, Laffan RE, Correia S, de Montaudouin X, Mesmer-
848 Dudons N, Freitas R, Culloty SC (2021) Latitudinal influence on gametogenesis
849 and host-parasite ecology in a marine bivalve model. Ecol. Evol. 11:7029-7041

850 Malham SK, Hutchinson TH, Longshaw M (2012) A review of the biology of European
851 cockles (*Cerastoderma* spp.). J. Mar. Biol. Assoc. U. K. 92:1563-1577

852 Maroso F, Casanova A, do Prado FD, Bouza C, Pardo BG, Blanco A, Hermida M,
853 Fernandez C, Vera M, Martinez P (2018) Species identification of two closely
854 exploited flatfish, turbot (*Scophthalmus maximus*) and brill (*Scophthalmus*
855 *rhombus*), using a ddRADseq genomic approach. Aquat. Conserv. 28:1253-1260

856 Maroso F, Perez de Gracia C, Iglesias D, Cao A, Diaz S, Villalba A, Vera M, Martinez
857 P (2019) A useful SNP panel to distinguish two cockle species, *Cerastoderma*
858 *edule* and *C. glaucum*, co-occurring in some European beds, and their putative
859 hybrids. Genes 10:760

860 Marquardt DW (1970).Generalized inverses, ridge regression, biased linear estimation
861 and nonlinear estimation. Technometrics 12:59.

862 Marquina D, Angel Fernandez-Alvarez F, Norena C (2015) Five new records and one
863 new species of Polycladida (Platyhelminthes) for the Cantabrian coast (North
864 Atlantic) of the Iberian Peninsula. J. Mar. Biol. Assoc. U. K. 95:311-322

865 Martinez L, Mendez J, Insua A, Arias-Perez A, Freire R (2013) Genetic diversity and
866 population differentiation in the cockle *Cerastoderma edule* estimated by
867 microsatellite markers. Helgol. Mar. Res. 67: 179-189

868 Martinez L, Freire R, Arias-Perez A, Mendez J, Insua A (2015) Patterns of genetic
869 variation across the distribution range of the cockle *Cerastoderma edule* inferred
870 from microsatellites and mitochondrial DNA. Mar. Biol. 162:1393-1406

871 Miller JM, Cullingham CI, Peery RM (2020) The influence of a priori grouping on
872 inference of genetic clusters: simulation study and literature review of the DAPC
873 method. Heredity 125:269-280

874 Nanton A, Arias-Perez A, Freire R, Fernandez-Perez J, Novoa S, Mendez J (2017)
875 Microsatellite variation in *Donax trunculus* from the Iberian Peninsula, with
876 particular attention to Galician estuaries (NW Spain). Estuar. Coast. Shelf Sci.
877 197:27-34

878 Nicolle A, Moitie R, Ogor J, Dumas F, Foveau A, Foucher E, Thiebaut E (2016)
879 Modelling larval dispersal of *Pecten maximus* in the English Channel: A tool for
880 the spatial management of the stocks. ICES J. Mar. Sci. 74:1812–1825.

881 Norris K, Bannister RCA, Walker PW (1998) Changes in the number of oystercatchers
882 *Haematopus ostralegus* wintering in the Burry Inlet in relation to the biomass of
883 cockles *Cerastoderma edule* and its commercial exploitation. J. Appl. Ecol.
884 35:75-85

885 Oksanen J. (2015) Multivariate analysis of ecological communities in R: vegan tutorial.
886 http://cc.oulu.fi/~jarioksa/opetus/metodi/vegantutor.pdf

887 Pampin M, Casanova A, Fernandez C, Blanco A, Hermida M, Vera M, Pardo BG,
888 Coimbra RM, Cao A, Iglesias D, Carballal MJ, Villalba A, Martinez P (2023)
889 Genetic markers associated with divergent selection against the parasite
890 *Marteilia cochillia* in common cockle (*Cerastoderma edule*) using
891 transcriptomics and population genomics data. Front. Mar. Sci. 10:1057206

892 Pardo BG, Fernandez C, Pampin M, Blanco A, Iglesias D, Cao A, Carballal MJ,
893 Villalba A, Martinez P (2022) Transcriptome characterization of the common

894 cockle (*Cerastoderma edule*) after exposure to a *Marteilia cochillia* outbreak.
895 Biorxiv doi10.1101/2022.10.18.512677

896 Parrondo M, Moran P, Ballenghien M, Acuna JL, Aguion A, Arrontes J, Chiss J, Cruz
897 T, Fernandes JN, Garcia-Florez L, Garcia-Vazquez E, Geiger KJ, Macho G,
898 Thiebaut E, Weidberg N, Jollivet D, Borrell YJ (2022) Chaotic genetic
899 patchiness in the highly valued atlantic stalked barnacle *Pollicipes pollicipes*
900 from the iberian peninsula: implications for fisheries management. Front. Mar.
901 Sci. 9:801780

902 Piñeira J, Quesada H, Rolan-Alvarez E, Caballero A (2008) Genetic discontinuity
903 associated with an environmentally induced barrier to gene exchange in the
904 marine snail *Littorina saxatilis*. Mar. Ecol. Prog. Ser. 357:175-184

905 Prive F, Luu K, Vilhjalmsson BJ, Blum MGB (2020) Performing highly efficient
906 genome scans for local adaptation with R Package pcadapt Version 4. Mol. Biol.
907 Evol. 37:2153-2154

908 Raj A, Stephens M, Pritchard JK (2014) fastSTRUCTURE: Variational inference of
909 population structure in large SNP data sets. Genetics 197:573-U207

910 Rellstab C, Gugerli F, Eckert AJ, Hancock AM, Holderegger R (2015) A practical guide
911 to environmental association analysis in landscape genomics. Mol. Ecol.
912 24:4348-4370

913 Richardson JL, Urban MC, Bolnick DI, Skelly DK (2014) Microgeographic adaptation
914 and the spatial scale of evolution. Trends Ecol. Evol. 29:165-176

915 Rohlf F (1993) NTSYS-pc. Numerical taxonomy and multivariate analysis system,
916 Version 2.1. Setauket, New York.

Rousset F (1997) Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. Genetics 145:1219-1228

Rousset F (2008) GENEPOP ' 007: a complete re-implementation of the GENEPOP software for Windows and Linux. Mol. Ecol. Resour. 8:103-106

Saco A, Suárez H, Novoa B, Figueras A (2023) A genomic and transcriptomic analysis of the C-type lectin gene family reveals highly expanded and diversified repertoires in bivalves. Mar. Drugs 21:254

Sa-Pinto A, Branco MS, Alexandrino PB, Fontaine MC, Baird SJE (2012) Barriers to gene flow in the marine environment: insights from two common intertidal limpet species of the Atlantic and Mediterranean. PLOS One 7:0050330

Sharples J, Simpson JH (2019) Shelf sea and shelf slope fronts. In: Reference Module in Earth Systems and Environmental Sciences. Vol. 1, pp 24-34. Elsevier.

Simpson JH, Pingree RD (1978). Shallow Sea Fronts Produced by Tidal Stirring. In: Bowman, M.J., Esaias, W.E. (eds) Oceanic Fronts in Coastal Processes. Springer, Berlin, Heidelberg.

Sokolov EP, Markert S, Hinzke T, Hirschfeld C, Becher D, Ponsuksili S, Sokolova IM (2019) Effects of hypoxia-reoxygenation stress on mitochondrial proteome and bioenergetics of the hypoxia-tolerant marine bivalve *Crassostrea gigas*. J. Proteomics 194:99-111

Suberg LA, Miller PI, Wynn RB (2019) On the use of satellite-derived frontal metrics in time series analyses of shelf-sea fronts, a study of the Celtic Sea. Deep Sea Res. Part I Oceanogr. Res. Pap. 149:103033

Theodosopoulos AN, Hund AK, Taylor SA (2019) Parasites and host species barriers in animal hybrid zones. Trends Ecol. Evol. 34:19-30

941 Thoma J, Stenitzer D, Grabherr R, Staudacher E (2022) Identification, characterization,
942 and expression of a beta-Galactosidase from *Arion* species (Mollusca).
943 Biomolecules 12:1578

944 Vendrami DLJ, De Noia M, Telesca L, Handal W, Charrier G, Boudry P, Eberhart-
945 Phillips L, Hoffman JI (2019) RAD sequencing sheds new light on the genetic
946 structure and local adaptation of European scallops and resolves their
947 demographic histories. Sci. Rep. 9:7455

948 Vera M, Carlsson J, El Carlsson J, Cross T, Lynch S, Kamermans P, Villalba A, Culloty
949 S, Martinez P (2016) Current genetic status, temporal stability and structure of
950 the remnant wild European flat oyster populations: conservation and restoring
951 implications. Mar. Biol. 163:239

952 Vera M, Maroso F, Wilmes SB, Hermida M, Blanco A, Fernandez C, Groves E,
953 Malham SK, Bouza C, Robins PE, Martinez P, Cockles C (2022) Genomic
954 survey of edible cockle (*Cerastoderma edule*) in the Northeast Atlantic: A
955 baseline for sustainable management of its wild resources. Evol. Appl. 15:262-
956 285

957 Vera M, Pardo BG, Cao A, Vilas R, Fernandez C, Blanco A, Gutierrez AP, Bean TP,
958 Houston RD, Villalba A, Martinez P (2019) Signatures of selection for
959 bonamiosis resistance in European flat oyster (*Ostrea edulis*): New genomic
960 tools for breeding programs and management of natural resources. Evol. Appl.
961 12:1781-1796

962 Vilas R, Vandamme SG, Vera M, Souza C, Maes GE, Volckaert FAM, Martinez P
963 (2015) A genome scan for candidate genes involved in the adaptation of turbot
964 (*Scophthalmus maximus*). Mar. Genom. 23: 77-86

965 Villalba A, Iglesias D, Ramilo A, Darriba S, Parada JM, No E, Abollo E, Molares J,
966 Carballal MJ (2014) Cockle *Cerastoderma edule* fishery collapse in the Ria de
967 Arousa (Galicia, NW Spain) associated with the protistan parasite *Marteilia*
968 *cochillia*. Dis. Aquat. Org. 109:55-80.

969 Wang M, Wang L, Ni D, Yi Q, Wang X, Jia Z, Song L (2018) The mRNA expression
970 profiles demonstrating versatile roles of glutathione S-transferase genes in the
971 mollusk *Chlamys farreri*. Invertebr. Surviv. J. 15:302-315

972 Waples RS (1998) Separating the wheat from the chaff: Patterns of genetic
973 differentiation in high gene flow species. J. Hered. 89:438-450

974 Woo S, Jeon HY, Kim SR, Yum S (2011) Differentially displayed genes with oxygen
975 depletion stress and transcriptional responses in the marine mussel, *Mytilus*
976 *galloprovincialis*. Comp. Biochem. Physiol. Part D Genomics Proteomics 6:
977 348-356

978

Figure legends:

Figure 1. Geographical distribution of the *Cerastoderma edule* beds analysed in the present study. Ocean bathymetry is shaded in blue. Summer sea surface ocean currents are schematically depicted with magenta coloured arrows. Tidal mixing fronts are indicated with purple dashed lines. 'UF' stands for Ushant Front, 'CSF' for Celtic Sea Front, and 'ISF' for Irish Sea Front. Location codes are shown in Table 1. Beds previously analysed by Vera et al. (2022) are marked with asterisks.

Figure 2. Population structure of *Cerastoderma edule* at different geographical scales using fastSTRUCTURE. Each vertical bar represents one individual, and the colour proportion of each bar represents the posterior probability of assignment of each individual to the different clusters (K) inferred by the program. The most likely K = 3 using the whole dataset (A), and K = 2 using the 57 shared divergent outliers between methodologies and K = 3 using the total 186 divergent outliers (B) for the SW British Isles are represented. Codes are shown on Table 1. Plots for all the K values tested for the different datasets are shown in Supplementary Figures.

Figure 3. Discriminant Analysis of Principal Components (DAPC) plots of *Cerastoderma edule* beds belonging to the SW British Isles. The weight of retained discriminant analysis (DA) and principal components selected are shown on left bottom box and right bottom box, respectively. Results using the complete whole dataset (A), the neutral dataset (B), the 35 shared divergent outliers between methodologies (C) and the 186 divergent outliers (D) are represented. Codes are shown on Table 1.

Figure 4. Discriminant Analysis of Principal Components (DAPC) plots of *Cerastoderma edule* beds belonging to Galicia. The weight of retained discriminant analysis (DA) and principal components selected are shown on left bottom box and

1003 right bottom box, respectively. Results using the complete dataset (A), neutral dataset

1004 (B) and 51 divergent outliers (C) are represented. Codes are shown on Table 1.

Table 1. *Cerastoderma edule* beds analysed in the present study. Location, sampling year, geographical coordinates, code, country, number of individuals collected (N initial) and analysed after quality filtering (N), observed heterozygosity (Ho), expected heterozygosity (He), inbreeding coefficient (F_{IS}) for all dataset and for polymorphic loci are shown. Locations in italics were previously analysed by Vera et al. (2022).

								All dataset			Polymorphic loci (MAF > 0.017)				
Location	**Year**	**Lat (deg N)**	**Lon (deg E)**	**Code**	**Country**	**N initial**	**N**	**Ho**	**He**	**F_{IS}**	**Polymorphic loci**	**% Polymorphic loci**	**Ho**	**He**	**F_{IS}**
Galway Connemara	2020	53.306	-9.846	IGC_20	Ireland	27	27	0.080	0.085	0.067	4205	45.5	0.166	0.179	0.074
Kerry- Feale	2020	52.488	-9.652	IKF_20	Ireland	30	30	0.080	0.087	0.080	4362	47.2	0.161	0.176	0.088
West Cork-Cockle Beach	2020	51.463	-9.744	IWC_20	Ireland	30	30	0.080	0.085	0.064	4294	46.4	0.163	0.175	0.070
Dundalk Bay-Annagassan	*2018*	*53.884*	*-6.341*	*IDA_18*	*Ireland*	*29*	*29*	0.074	0.080	0.082	*2993*	32.4	*0.161*	*0.177*	0.090
Dundalk Bay-Cooley	*2018*	*53.996*	*-6.287*	*IDC_18*	*Ireland*	*22*	*22*	0.077	0.083	0.081	*3493*	37.8	*0.173*	*0.190*	0.087
Dee Estuary	*2017*	*53.343*	*-3.174*	*WDE_17*	*Wales*	*30*	*28*	0.071	0.076	0.074	*2561*	27.7	*0.167*	*0.182*	0.084
Burry	*2017*	*51.643*	*-4.166*	*WBY_17*	*Wales*	*30*	*30*	0.073	0.080	0.091	*3529*	38.2	*0.148*	*0.165*	0.101
Camel Estuary (Cornwall)	2020	50.531	-4.930	ECE_20	England	24	24	0.073	0.081	0.105	3667	39.6	0.159	0.180	0.114
Somme Bay	*2017*	*50.201*	*1.627*	*FBS_17*	*France*	*30*	*30*	0.071	0.080	0.111	*3438*	37.2	*0.147*	*0.167*	0.119
Baie des Veys (Brévands)	2019	49.365	-1.150	FBV_19	France	26	26	0.079	0.081	0.022	3579	38.7	0.169	0.174	0.030
Gouville sur mer	2019	49.105	-1.612	FGO_19	France	23	23	0.077	0.082	0.059	4055	43.8	0.163	0.174	0.063
Arcachon Bay	*2017*	*44.580*	*-1.238*	*FAR_17*	*France*	*30*	*30*	0.074	0.083	0.111	*4335*	46.9	*0.140*	*0.159*	0.120
O Barqueiro	2017	43.722	-7.701	SBA_17	Spain	30	30	0.076	0.085	0.107	4595	49.7	0.140	0.158	0.115
Miño	2017	43.361	-8.206	SMI_17	Spain	30	30	0.073	0.081	0.102	4159	45.0	0.140	0.157	0.110
Anllóns	2017	43.220	-8.943	SAN_17	Spain	30	29	0.073	0.081	0.101	3526	38.1	0.143	0.161	0.110
Ría de Noia	*2017*	*42.790*	*-8.923*	*SNO_17*	*Spain*	*30*	*30*	0.078	0.087	0.099	*4885*	52.8	*0.139*	*0.156*	0.107
Lombos do Ulla	*2017*	*42.629*	*-8.775*	*SLO_17*	*Spain*	*30*	*30*	0.075	0.085	0.113	*4602*	49.8	*0.137*	*0.156*	0.120
Sarrido	2017	42.507	-8.826	SSA_17	Spain	30	30	0.074	0.083	0.103	4317	46.7	0.140	0.158	0.112
Vilanova	2017	42.561	-8.831	SVI_17	Spain	27	25	0.072	0.081	0.112	3175	34.3	0.153	0.173	0.118
Campelo	2017	42.421	-8.685	SCA_17	Spain	30	30	0.073	0.082	0.115	4287	46.3	0.140	0.159	0.121
Moaña	2017	42.286	-8.730	SMO_17	Spain	20	19	0.070	0.077	0.088	2361	25.5	0.181	0.200	0.093
Baiona	2018	42.117	-8.822	SBI_18	Spain	17	17	0.075	0.082	0.092	3609	39.0	0.174	0.192	0.097

1009 Table 2. Results of the redundancy analysis (RDA) on the SW British Isles region of *Cerastoderma edule*. Only variables included by the
1010 forward selection model are shown.

1011

SW British Isles			**Complete dataset**		**Neutral dataset**		**186 total outlier dataset**	
Model	**Season**	**Variable**	**P-value**	**Adjusted R^2**	**P-value**	**Adjusted R^2**	**P-value**	**Adjusted R^2**
All seascape variables	**Reproductive period**	Latitude	-	0.102	-	0.098	0.005	0.414
		Longitude	0.001		0.001		0.001	
		SBS	-		–		0.051	
		BSS	0.004		0.001		-	
	Winter	Latitude	0.001	0.087	0.001	0.080	0.009	0.423
		Longitude	0.016		0.015		0.001	
		SBS	-		-		0.030	
	Summer	Latitude	-	0.102	-	0.098	0.012	0.410
		Longitude	0.002		0.003		0.001	
		SBS	-		-		0.055	
		BSS	0.001		0.001		-	
Only abiotic variables	**Reproductive period**	SST	-	0.053	-	0.053	0.013	0.319
		SBT	-		-		0.001	
		BSS	0.002		0.002		-	
		NPP	-		-		0.027	
	Winter	SST	0.003	0.067	-	0.069	-	0.202
		SSS	-		-		0.041	
		BSS	-		0.010		-	

		NPP	0.017		0.008		0.019	
	Summer	SST	0.011	0.086	0.014	0.083	0.001	0.355
		SBT	-		-		0.009	
		SBS	-		-		0.083	
		BSS	0.001		0.001		-	

Galicia			**Complete dataset**		**Neutral dataset**		**51 total outlier dataset**	
Model	**Season**	**Variable**	**P-value**	**Adjusted R^2**	**P-value**	**Adjusted R^2**	**P-value**	**Adjusted R^2**
All seascape variables	**Reproductive period**	Latitude	-	-	-	-	0.012	0.124
	Winter	Latitude	-	-	-	-	0.011	0.124
	Summer	Latitude	-	-	-	-	0.008	0.124
Only abiotic variables	**Reproductive period**	-	-	-	-	-	-	-
	Winter	BSS	-	-	-	-	0.045	0.083
	Summer	-	-	-	-	-	-	-

Adjusted R^2 and P-value associated to each variable of its selection stage. SST: Sea Surface Temperature; SBT: Sea Bottom Temperature; SSS: Sea Surface Salinity; SBS: Sea Bottom Salinity; BSS: Bottom Shear Stress; NPP: Net Primary Production.

Table 3. Distribution of divergent outliers and markers correlated with environmental variables in the SW British Isles and Galicia across the *Cerastoderma edule* genome (version 4.0).

		Outlier loci (divergent selection)			
Mega-scaffold (chromosome)	Chromosome length (bp)	British Isles	Galicia	Shared	Markers correlated envionmental variables*
C1	64,609,245	21	11	3	5 (3)
C2	56,319,168	14	2		7 (3)
C3	55,987,847	22	3	1	6 (5)
C4	52,087,795	18	5	2	7 (6)
C5	50,828,891	11	1	1	3 (2)
C6	40,237,005	13	3	2	4 (1)
C7	39,934,596	2	1		
C8	39,684,391	9			2 (1)
C9	39,070,162	11	3		2 (1)
C10	38,264,924	14	8	2	1
C11	38,197,540	2			1
C12	36,327,582	6	1		1 (1)
C13	35,955,507	10			5 (2)
C14	33,816,358	5			
C15	31,726,440	3	1	1	3 (1)
C16	31,510,408	10	8	1	2 (2)
C17	26,587,828	4			2 (1)
C18	22,603,465	1	1	1	
C19	21,711,631	4	1	1	1
Other scaffolds		6	2		2 (1)
Total		186	51	15	54 (30)

* Only detected in SW British Isles; in parentheses those markers also identified as outliers for divergent selection

Fig.1

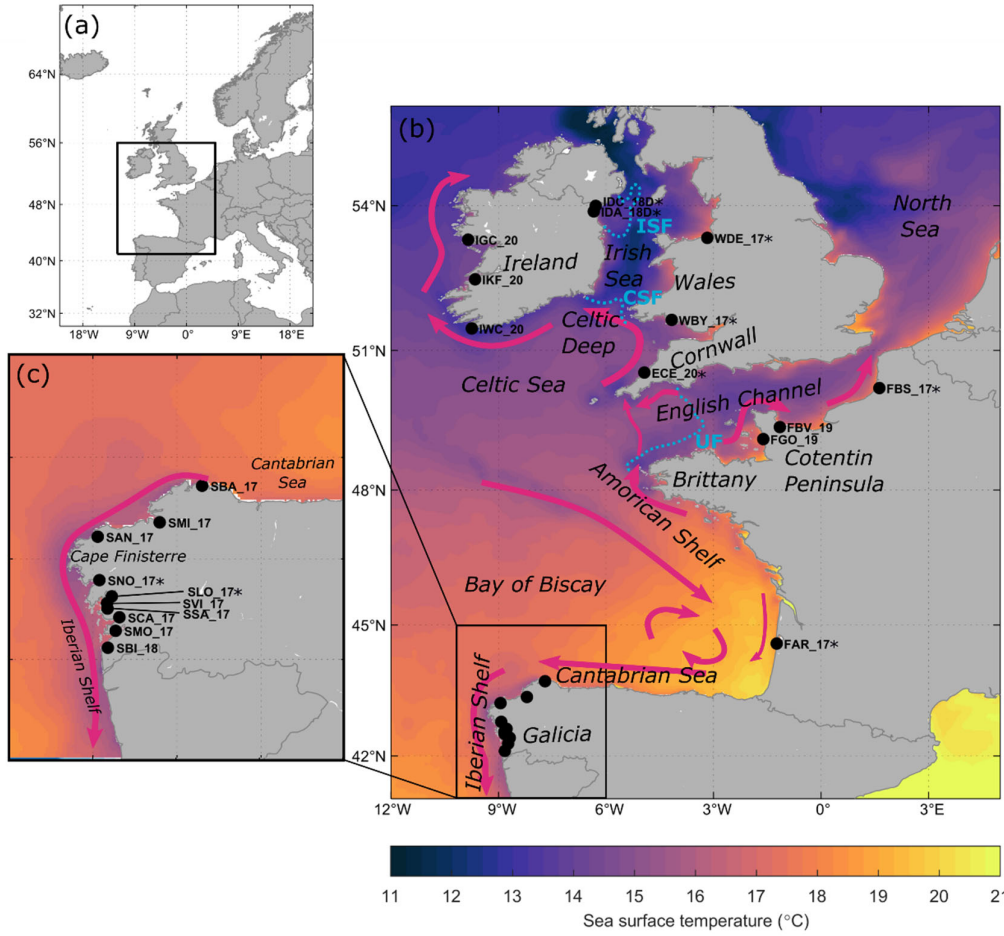

(a)
64°N
56°N
48°N
40°N
32°N
18°W
9°W
0°
9°E
18°E
(b)
54°N
51°N
48°N
45°N
42°N
12°W
9°W
6°W
3°W
0°
3°E
IDL_18D*
IDA_18D*
ISF
IGC_20
WDE_17*
Ireland
Irish Sea
Wales
North Sea
IKF_20
CSF
IWC_20
Celtic Deep
WBY_17*
Cornwall
ECE_20*
Celtic Sea
English Channel
FBS_17*
FBV_19
FGO_19
UF
Cotentin Peninsula
Brittany
Armorican Shelf
Bay of Biscay
FAR_17*
Cantabrian Sea
Iberian Shelf
Galicia
(c)
Cantabrian Sea
SBA_17
SMI_17
SAN_17
Cape Finisterre
SNO_17*
SLO_17*
SVI_17
SSA_17
SCA_17
SMO_17
SBI_18
Iberian Shelf
11
12
13
14
15
16
17
18
19
20
21
Sea surface temperature (°C)

Fig.2

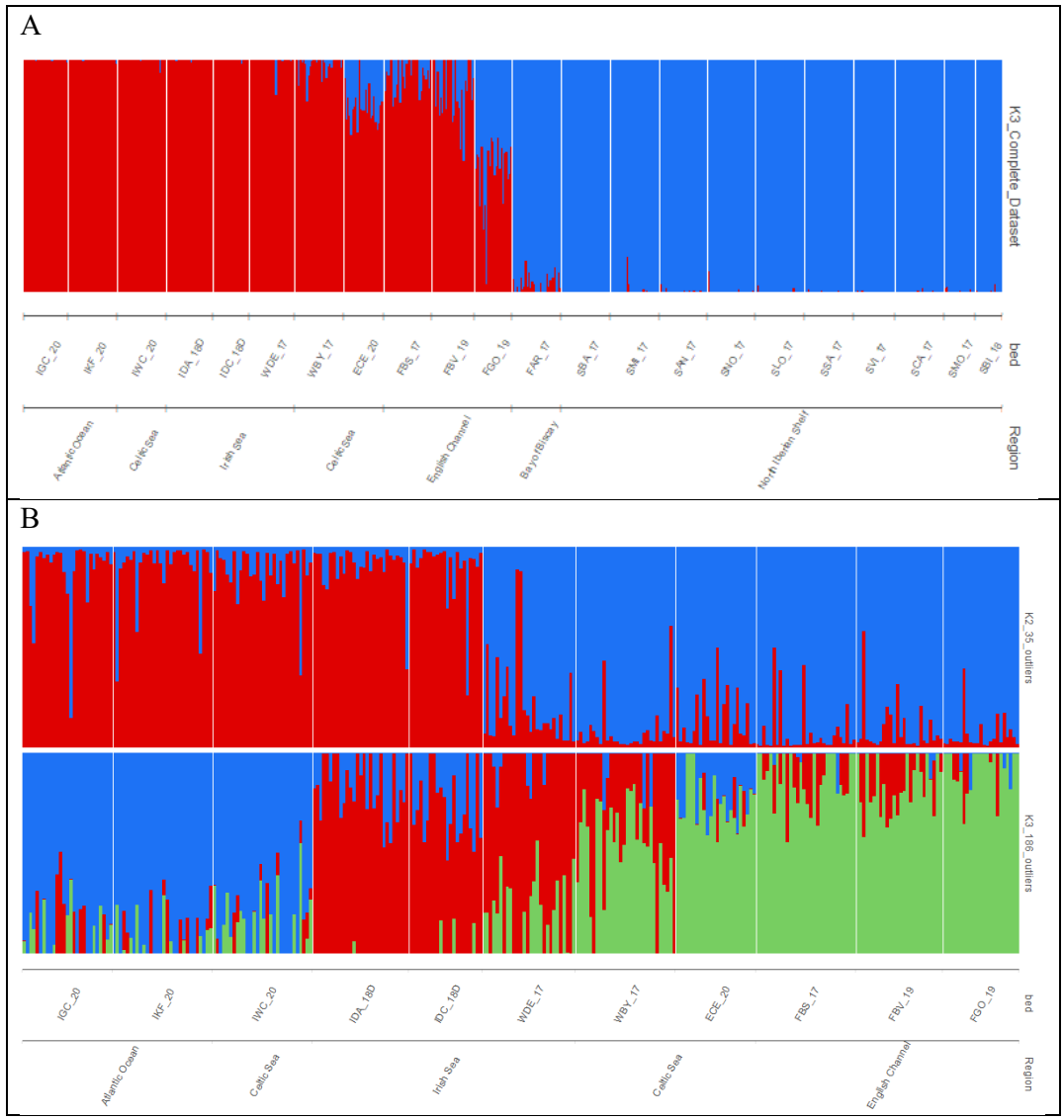
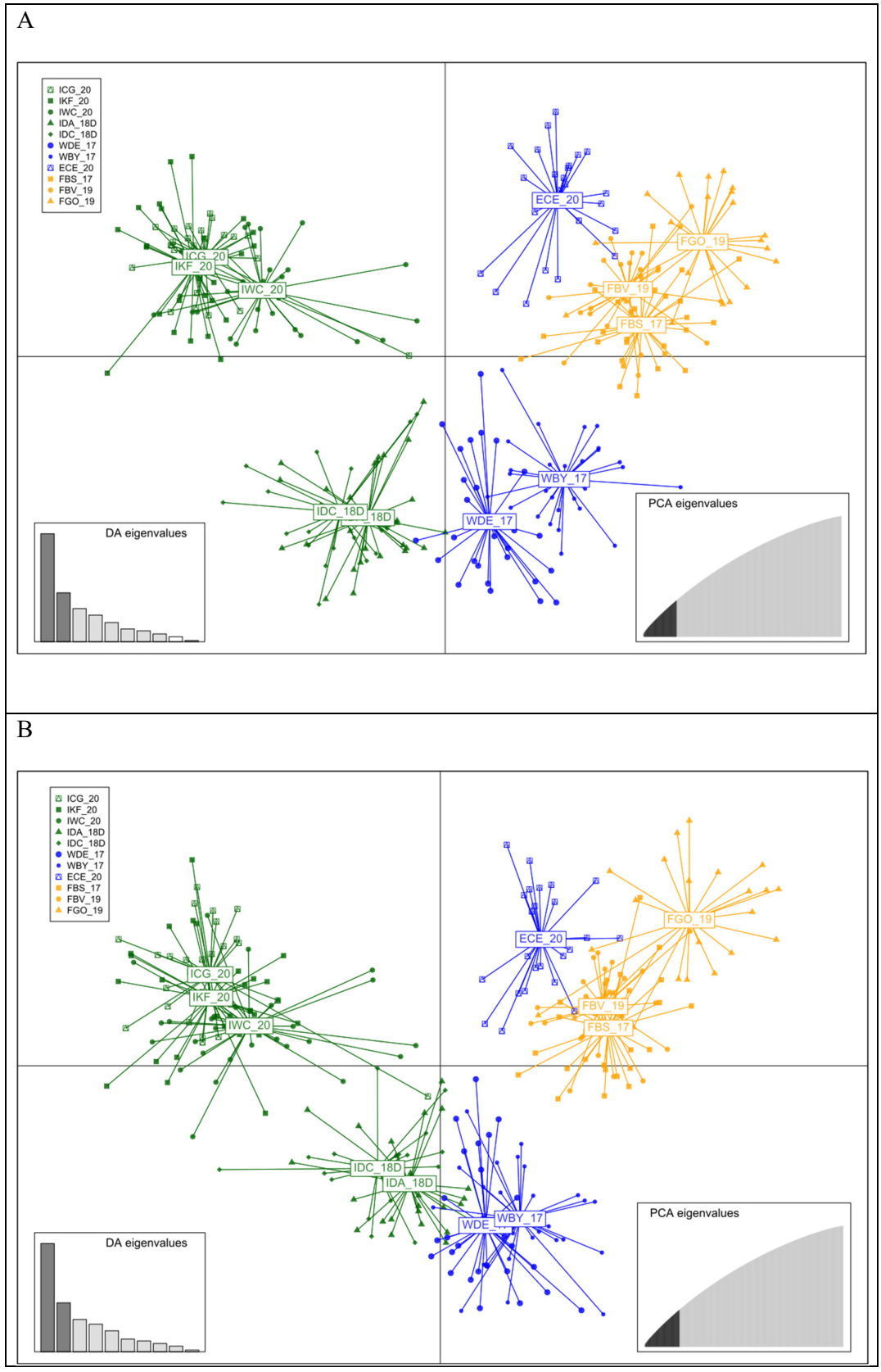
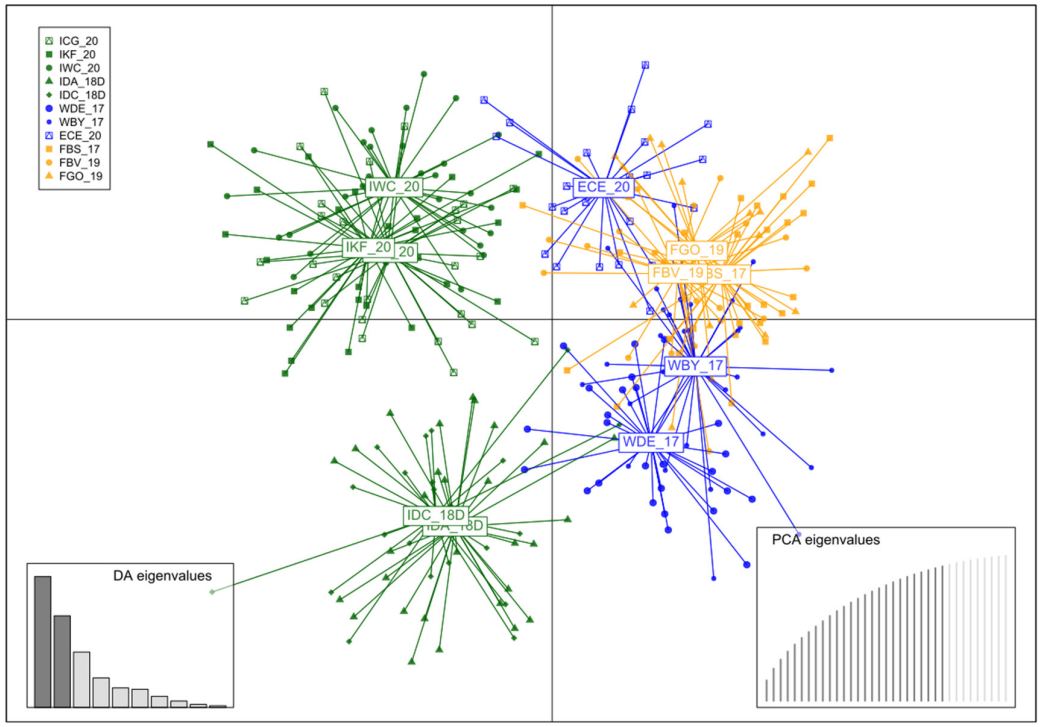

A
K3_Complete_Dataset
bed
Region
IGC_20
IKF_20
IWC_20
IDA_18D
IDC_18D
WDE_17
WBY_17
ECE_20
FBS_17
FBV_19
FGO_19
FAR_17
SBA_17
SMI_17
SWI_17
SKO_17
SLO_17
SSA_17
SVI_17
SCA_17
SMO_17
SBI_16
Atlantic Ocean
Celtic Sea
Irish Sea
Celtic Sea
English Channel
Bay of Biscay
North Iberian Shelf
B
K2_35_outliers
K3_186_outliers
bed
Region
IGC_20
IKF_20
IWC_20
IDA_18D
IDC_18D
WDE_17
WBY_17
ECE_20
FBS_17
FBV_19
FGO_19
Atlantic Ocean
Celtic Sea
Irish Sea
Celtic Sea
English Channel

Fig.3


A
ICG_20
IKF_20
IWC_20
IDA_18D
IDC_18D
WDE_17
WBY_17
ECE_20
FBS_17
FBV_19
FGO_19
ICG_20
IKF_20
IWC_20
ECE_20
FGO_19
FBV_19
FBS_17
IDC_18D
18D
WBY_17
WDE_17
DA eigenvalues
PCA eigenvalues
B
ICG_20
IKF_20
IWC_20
IDA_18D
IDC_18D
WDE_17
WBY_17
ECE_20
FBS_17
FBV_19
FGO_19
ICG_20
IKF_20
IWC_20
ECE_20
FGO_19
FBV_19
FBS_17
IDC_18D
IDA_18D
WDE
WBY_17
DA eigenvalues
PCA eigenvalues

C


ICG_20
IKF_20
IWC_20
IDA_18D
IDC_18D
WDE_17
WBY_17
ECE_20
FBS_17
FBV_19
FGO_19
IWC_20
ECE_20
IKF_20
FGO_19
FBV_19
WBY_17
WDE_17
IDC_18D
DA eigenvalues
PCA eigenvalues

D

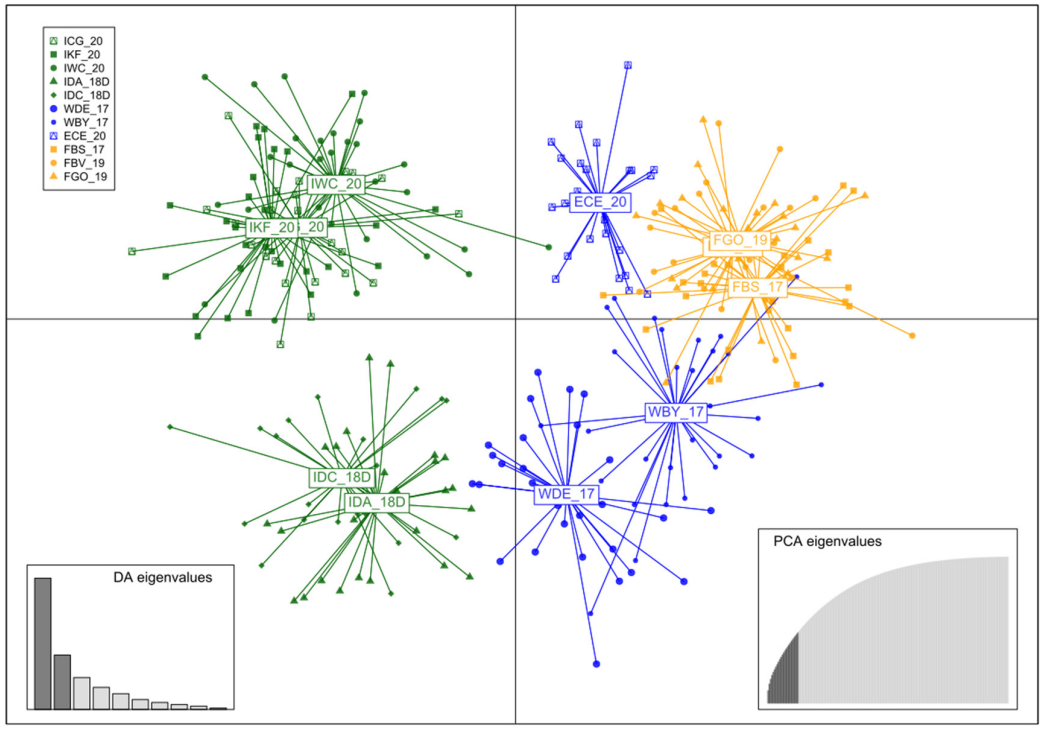
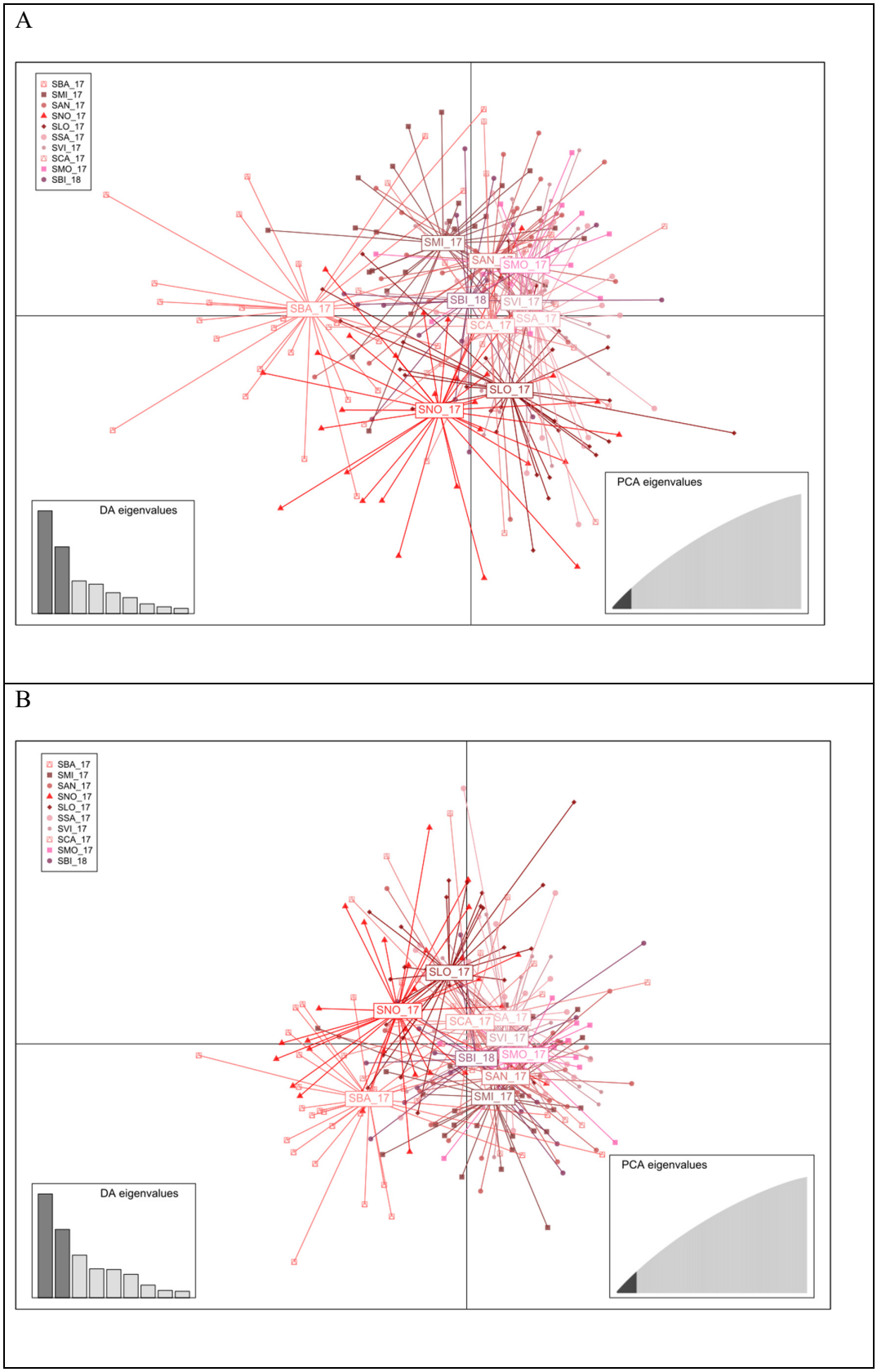
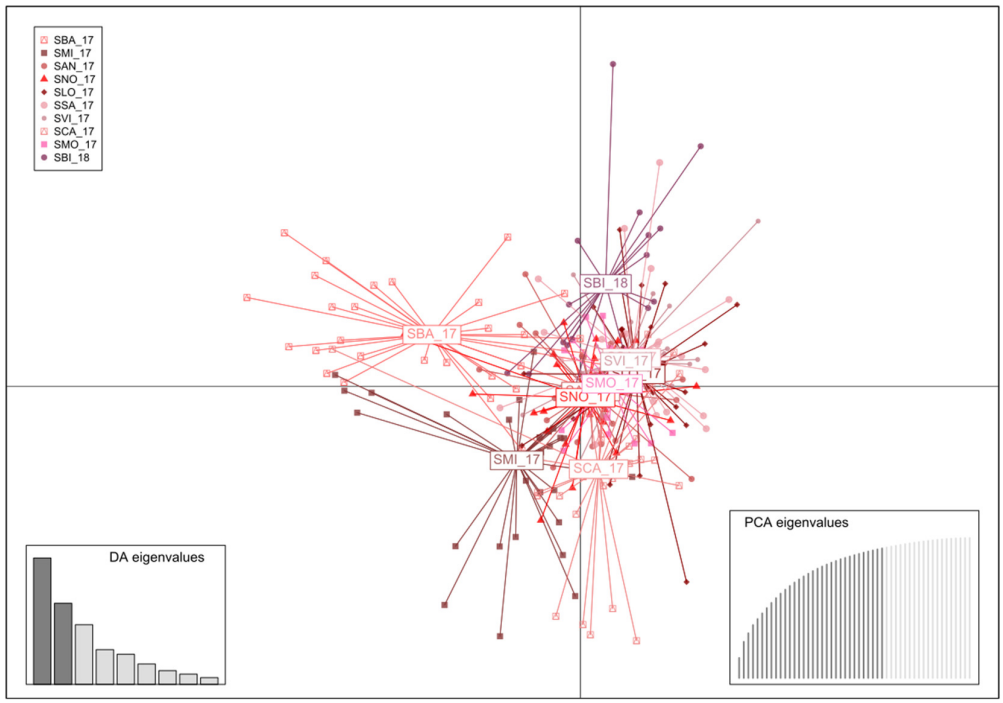

ICG_20
IKF_20
IWC_20
IDA_18D
IDC_18D
WDE_17
WBY_17
ECE_20
FBS_17
FBV_19
FGO_19
IWC_20
ECE_20
FGO_19
FBS_17
WBY_17
IDC_18D
IDA_18D
WDE_17
DA eigenvalues
PCA eigenvalues

Fig. 4


A
SBA_17
SMI_17
SAN_17
SNO_17
SLO_17
SSA_17
SVI_17
SCA_17
SMO_17
SBI_18
SMI_17
SAN
SMO_17
SBI_18
SVI_17
SBA_17
SCA_17
SSA_17
SLO_17
SNO_17
DA eigenvalues
PCA eigenvalues
B
SBA_17
SMI_17
SAN_17
SNO_17
SLO_17
SSA_17
SVI_17
SCA_17
SMO_17
SBI_18
SLO_17
SNO_17
SCA_17
SVI_17
SBI_18
SMO_17
SAN_17
SBA_17
SMI_17
DA eigenvalues
PCA eigenvalues

C


SBA_17
SMI_17
SAN_17
SNO_17
SLO_17
SSA_17
SVI_17
SCA_17
SMO_17
SBI_18
SBA_17
SBI_18
SVI_17
SMO_17
SNO_17
SMI_17
SCA_17
DA eigenvalues
PCA eigenvalues