Oceanographic currents, geographic patterns and local environment contribute to neutral and adaptive genetic structure in two intertidal marine gastropods with contrasting life-histories

Melanie Prentice[1], Sophie Ward[2], Peter Robins[2], Paul Shaw[3], Hannah Earp[3], Colin Sauze[3], Ilze Skujina[3], Stuart Jenkins[2], and Joseph Ironside[3]

[1]The University of Maine
[2]Bangor University
[3]Aberystwyth University

February 22, 2024

Abstract

Global climate change is exposing intertidal organisms to increasing air and sea temperatures and changing ocean currents, affecting their ability to disperse, survive and reproduce, and resulting in shifts in their distribution and abundance. Improved understanding of these shifts requires characterization of population structure and local adaptation. We estimate the drivers of population structure in two intertidal gastropod species with contrasting life histories by assessing neutral and adaptive population structure and performing redundancy analyses in a seascape genomics framework. We show putative adaptive divergence between populations of the broadcast spawning topshell, Steromphala umbilicalis, despite high rates of neutral gene flow. This adaptive structure was best explained by geographic structure, separating sites in Wales from all other British and Irish sites. Larval dispersal, estimated from biophysical models, was also identified as a minor component explaining genetic connectivity in this species. For the direct developing dogwhelk, Nucella lapillus¸ neutral population structure was best explained by air and sea surface temperatures while putative adaptive population structure showed a greater influence of wave exposure. These results support the expected relationship between neutral population structure and reproductive mode (i.e., greater population structure in the direct developing N. lapillus compared with a lack of structure in the broadcast spawning S. umbilicalis) and highlight the interactive effects of geographic structure, larval dispersal and local environment on gene flow and adaptation of intertidal marine organisms.

Oceanographic currents, geographic patterns and local environment contribute to neutral and adaptive genetic structure in two intertidal marine gastropods with contrasting life-histories

Running title: seascape genomics of intertidal invertebrates

Melanie B Prentice[a] 11School of Food and Agriculture, University of Maine, Orono, Maine, USA[2]Department of Ichthyology & Fisheries Science, Rhodes University, Makhanda, South Africa[3]School of Natural and Environmental Science, The Dove Marine Laboratory, Newcastle University, North Shields, England**, Sophie L Ward[b], Peter E Robins[b], Paul W Shaw[a2], Hannah S Earp[a3], Colin Sauze[a], Ilze Skujina[a], Stuart R Jenkins[b], Joseph E Ironside[a*]**

[a]Institute of Biological, Environmental, and Rural Sciences, Aberystwyth University, Aberystwyth, Wales

[b]School of Ocean Sciences, Bangor University, Menai Bridge, Anglesey, Wales

*corresponding author; jei@aber.ac.uk

ABSTRACT

Global climate change is exposing intertidal organisms to increasing air and sea temperatures and changing ocean currents, affecting their ability to disperse, survive and reproduce, and resulting in shifts in their distribution and abundance. Improved understanding of these shifts requires characterization of population structure and local adaptation. We estimate the drivers of population structure in two intertidal gastropod species with contrasting life histories by assessing neutral and adaptive population structure and performing redundancy analyses in a seascape genomics framework. We show putative adaptive divergence between populations of the broadcast spawning topshell, *Steromphala umbilicalis* , despite high rates of neutral gene flow. This adaptive structure was best explained by geographic structure, separating sites in Wales from all other British and Irish sites. Larval dispersal, estimated from biophysical models, was also identified as a minor component explaining genetic connectivity in this species. For the direct developing dogwhelk, *Nucella lapillus* ¸ neutral population structure was best explained by air and sea surface temperatures while putative adaptive population structure showed a greater influence of wave exposure. These results support the expected relationship between neutral population structure and reproductive mode (i.e., greater population structure in the direct developing *N. lapillus* compared with a lack of structure in the broadcast spawning *S. umbilicalis*) and highlight the interactive effects of geographic structure, larval dispersal and local environment on gene flow and adaptation of intertidal marine organisms.

Keywords: seascape genomics, molluscs, intertidal, connectivity, natural selection, adaptation

INTRODUCTION

Understanding genetic diversity and connectivity between populations of marine species is critical to inform the conservation and management of the marine environment (Palumbi 2004; Baco et al. 2016; Xuereb et al. 2019). Historically, marine populations have been assumed to be demographically open, owing to their large population sizes, high larval dispersal potential and the perceived lack of physical barriers to gene flow in the marine environment (Palumbi 1994; Waples 1998). However, several studies have shown distinct patterns of population structure in marine species, even at small spatial scales (e.g., Hoffman et al. 2011a; Benestan et al. 2015; D'Aloia et al. 2015; DiBattista et al. 2017; Coscia et al. 2020), supporting the prospect of a seascape fragmented not only by geographical barriers, but by environmental gradients (e.g., ocean currents, tidal mixing fronts, temperature) that can act as either conduits or barriers to dispersal (Galindo et al. 2006; Treml et al. 2015; Benestan et al. 2016).

The dispersal potential of marine invertebrates and hence gene flow within and among populations is driven not only by the environment but also by their life history strategy (Shanks 2009; Selkoe & Toonen 2011). The majority of benthic marine invertebrates have a complex life cycle with a planktonic larval phase (Thorson 1950). Larvae, especially those with long planktonic life spans, are expected to disperse widely, leading to high levels of gene flow. The relationship between pelagic larval duration (PLD) and genetic connectivity, however, has been shown to be weak in some cases (Weersing & Toonen 2009). Even in taxa without larval stages (i.e., those with direct developing juveniles), there are contradictory patterns. Whilst low levels of connectivity and hence gene flow might be expected in such species (e.g., Sherman et al. 2008; Hoffman et al. 2011b; Pascual et al. 2017), molecular studies have provided an increasing number of exceptions to this assumption (e.g., Edmands & Potts 1997; Kyle & Boulding 2000; Ayre & Hughes 2000; Richards 2007).

Beyond estimating genetic connectivity, genome-wide datasets facilitate the characterization of the adaptive potential of populations. Further, when aligned with spatial and environmental data, we can identify the variables underlying the organization of adaptive genetic structure (i.e., the seascape genomics framework; Selkoe et al. 2010, 2016; White et al. 2010; Benestan et al. 2016; Bernatchez et al. 2019). Given that marine species are generally characterised by large effective population sizes (*Ne*), and the efficacy of selection increases with increasing *Ne* (via a reduction in genetic drift), marine populations should be particularly well suited to exhibit local adaptation (Allendorf et al. 2010; Gagnaire et al. 2015). The theoretical expectation that selection is counteracted by the homogenizing effects of gene flow (Lenormand 2002) suggests that responses to selection may differ between species with different propensities for gene flow. Specifically,

species with more restricted gene flow (i.e., direct developing species) may demonstrate stronger effects of selection than species that disperse more widely (i.e., broadcast spawning species), in which locally adapted populations will be swamped by the immigration of maladapted genotypes (e.g., Yamada 1989). Recent evidence, however, suggests that local adaptations can persist in marine populations despite ongoing gene flow (Tigano & Friesen 2016; Hoey & Pinsky 2018; Sandoval-Castillo et al. 2018; Teske et al. 2019).

Here, we estimate the neutral and adaptive genetic structure of two similarly distributed rocky intertidal gastropods (*Nucella lapillus* and *Steromphala umbilicalis*) in the Irish Sea using restriction site-associated DNA (RADseq) datasets. *S. umbilicalis*is a broadcast spawning gastropod mollusc inhabiting mid- to low-intertidal zones from Morocco (Southward et al. 1995) to north-west Scotland (Mieszkowska et al. 2013). It produces lecithotrophic larvae with a maximum larval duration of 7-9 days (Keith et al. 2011). *N. lapillus* is a gastropod mollusc of the low intertidal zone within the north Atlantic (Wares & Cunningham 2001). It reproduces by ovo-viviparity; fertilization is internal, egg capsules are laid on rocky intertidal surfaces, and juvenile snails emerge after 8-12 weeks (Strathmann 1987).

The combination of limited adult movement (e.g., up to 30 m; Castle & Emery 1981) and direct development of juveniles leads to an expectation of low gene flow and high differentiation in *N. lapillus* . Although there has been some support for this (e.g., Rolán et al. 2004; Bell & Okamura 2005), in many cases, evidence of higher gene flow has emerged (e.g., Colson & Hughes 2007; Bell 2008). For example, Colson & Hughes (2004) demonstrated levels of genetic differentiation in the British Isles that were comparable to that of species with short-lived pelagic larval reproduction. These results support rapid and common medium-distance movements in *N. lapillus* (10-100 km), probably through the passive transport of egg masses or small juveniles via rafting. Further, clinal genetic differentiation over environmental gradients, and genetic differentiation between morphs support the contribution of heritable adaptive genetic differentiation to ecotypes of *N. lapillus* (Kirby 2000; Rolán et al. 2004; Guerra-Varela et al. 2009; Carro et al. 2019). While less is known about the population structure of *S. umbilicalis* , Wort et al. (2019) showed no evidence of genetic differentiation between populations separated by an ˜230 km habitat gap, suggesting that currents, PLD and spawning season are more important than distance in determining genetic differentiation in this species.

The rocky habitat required by our model species is linear and contains long gaps formed by stretches of soft sediment or open sea. This allows us to compare connectivity of a direct developing vs. a larval dispersing species across habitat gaps using a combination of neutral and adaptive genetic markers. With regard to neutral genetic variation, we predict that the larval dispersing *S. umbilicalis* will show greater connectivity across large habitat gaps, whereas the direct developing *N. lapillus*will show finer scale local structure. With regard to adaptive genetic variation, we hypothesize that selection will overcome the homogenizing effects of gene flow, resulting in more complex patterns of population structure associated with the key environmental variables driving selection in each species.

In addition to comparisons of genetic connectivity, we also estimate the relative contribution of physical and environmental variables (spatial structure, oceanography, climate) on patterns of neutral and adaptive genetic structure in a seascape genomics framework. Provided that oceanographic features are well represented by hydrodynamic model simulations, their influence on dispersal and thus connectivity can be estimated using larval dispersal algorithms that incorporate simulated ocean flows with species biological traits (e.g., Robins et al. 2013). Such methods can also be used to investigate stepping-stone dispersal over successive generations (e.g., Giménez et al. 2020a). In addition to oceanography, we investigate the influence of sea surface temperature (SST) and air temperature (AT) on population structure, both of which have been related to the distribution and abundance of our model species (Southward & Crisp 1954; Kendall et al. 1987) and are important selective agents driving population divergence in marine invertebrates (Sanford & Kelly 2011). We also include wave exposure in our environmental dataset, as it has been shown to affect the abundance of*S. umbilicalis* (Ballantine 1961) and underlies the morphological variation delineating ecotypes of *N. lapillus* (Guerra-Varela et al. 2009; Carro et al. 2019).

METHODS

Study Sites & Sampling

We selected sites to maximize the range of climatic variation within our study area (Irish Sea), and included sites separated by potential barriers to dispersal and gene flow (e.g., divergent residual currents and habitat gaps). We surveyed 120 sites consisting of either natural or artificial rocky substrates along the eastern coast of the Republic of Ireland and Northern Ireland, and the western coast of Wales, England and Scotland between June 2018 and June 2019 for the abundance of our two focal species (Fig. 1A). We focused on areas lacking suitable habitat that may represent large-scale barriers to dispersal. Two large stretches of coastline devoid of populations of one or both species were identified, representing putative dispersal barriers. The first of these was the eastern coastline of Ireland between Coliemore and Rosslare in which, of 25 sites surveyed, none contained *S. umbilicalis* and 17 (68%) contained *N. lapillus* . The northern coastline of Wales, east of Llanddulas to the southwestern coastline of Scotland at Brighouse Bay formed the second coastal gap in which, of 16 sites surveyed, two (12.5%) contained *S. umbilicalis* (with only one individual found at one of these sites) and five (31.3%) contained*N. lapillus* . The main channel of the Irish Sea was considered the third "habitat gap" separating Welsh, English, and Scottish sites from those in Ireland and Northern Ireland. Twelve sites were selected for genetic sampling, including sites at either end of each habitat gap where thriving populations of both species were found (Fig. 1). From each site, 20 adult individuals of each species were sampled resulting in 240 individuals per species. Tissues were frozen following collection, and subsequently preserved in 99% ethanol prior to DNA extraction.

Sequencing

DNA was extracted using a phenol-chloroform protocol (Appendix A). Library preparations and sequencing was conducted by Data2Bio LLC (Ames, Iowa), using the method outlined in Ott et al. (2017). Paired-end 150 bp sequencing was performed on an Illumina HiSeq X instrument.

Genomic Datasets & Bioinformatics

The quality of the demultiplexed raw data for each sample was checked using FastQC (v0.11.8; Andrews 2010) and summarized with MultiQC (v1.8; Ewels et al. 2016; Table S1). Low quality sequences were discarded, and the remainder were trimmed to uniform length (140bp) using the*process_ radtags* program in STACKS (v2.5; Catchen et al. 2013). The quality of the filtered datasets was then re-assessed using FastQC and MultiQC (Table S2). Identification and genotyping of polymorphic SNPs was conducted using the STACKS pipeline. We optimized the key parameters M, n and m, following the protocol of Paris et al. (2017; Appendix B). These parameters resulted in final datasets containing 2,494,940 and 2,325,766 polymorphic SNPs in *S. umbilicalis* and*N. lapillus* , respectively. The resulting datasets were filtered to remove SNPs and samples of low quality and/or with high proportions of missing data (Table 1; Appendix C), and file formats were converted using PGDSpider (v2.1.1.5; Lischer & Excoffier 2012) for all downstream analyses.

Environmental Dataset

As temperature influences intertidal organisms at all life-history stages, we calculated minimum, maximum, and average SST and AT in winter (December-February) and summer (June-September) months between 2008-2018. We included September in the summer calculations to encompass the entire potential larval period of *S. umbilicalis* . We used seasonal averages for the decade prior to and including the year of sampling so that the temporal resolution of the environmental data encompassed multiple generations of selective pressure (Flanagan et al. 2018). SST data was obtained as monthly means per year at a resolution of ~11 x 7.5 km^2 from the EU Copernicus Marine Service Information Atlantic-European North West Shelf Ocean Physics Reanalysis dataset (REF: CMEMS-NWS-PUM-004-009). AT data was obtained as temperature 2 m from the sea surface, at a resolution of 9 km^2 from the EU Copernicus Climate Change Service ERA5-Land dataset (DOI: 10.24381/cds.68d2bb30).

We also incorporated descriptors of historical climate using five bioclimatic variables obtained from the WorldClim dataset (Fick & Hijmans 2017): maximum AT of the warmest month, minimum AT of the coldest month, AT annual range, mean AT of the warmest quarter and mean AT of the coldest quarter.

These variables represent historic 30-year monthly averages from 1970-2000 at a resolution of 5 km^2.

Wave exposure data was taken from the UK dataset of Burrows et al. (2008), who developed a grid-based model to calculate indices of wave fetch, a parameter that describes the distance to the nearest land cell.

Most available records of environmental variables were either "land-based" or "sea-based" with few encompassing the intertidal environment. To address this, we took averages of all recorded temperature values within a 10 km radius of each sampling site. For site/environmental variable combinations where no data was available within this radius, we used the nearest data cell to characterize the site. As the exposure dataset is high-resolution coastal data, we were able to obtain more precise estimates for each site and retained the average wave fetch value of the nearest cell only. We used average wave fetch (representing the average of the summed fetch values for the cell and its immediate neighbours) to incorporate some variability in the local degree of exposure at each site. All environmental data manipulations were conducted in ArcGIS (v10.7.1; ESRI 2011).

We conducted a principal component analysis (PCA) in R (v4.0.3; R Core Team 2020) to reduce the set of 18 environmental variables into fewer, orthogonal PCs that describe environmental variation at our study sites. We retained PCs that had eigenvalues >1 and assessed the loadings of environmental variables onto each retained component. We used a cut-off value of 0.32 to attribute variables to each PC, as this threshold represents environmental variables that have at least 10% of their variance explained by the PC (Dormann et al. 2013). We then calculated the standardized scores of each retained PC for subsequent analyses.

Detection of Outlier Loci

Prior to assessing population structure and adaptation, we identified outlier loci (i.e., loci putatively under the influence of selection) and divided our datasets into (1) putatively adaptive and (2) putatively neutral markers for each species. To detect outlier loci, we used four population differentiation (PD) methods and two environmental association (EA) methods that implement different methodologies (Table 2, Appendix D).

The use of both PD and EA methods allows for the conservative detection of outlier loci by comparing complementary methods with different assumptions and biases. PD methods use measures of differentiation (F_{ST} or in the case of ordination approaches, multivariate distances) to identify outlier loci, where extreme levels of differentiation (i.e., those outside of neutral expectations) are likely candidates for selection (Liggins et al. 2019). EA methods use correlations between environmental variables and allele frequencies to identify loci putatively under selection (Rellstab et al. 2015). While both methods are prone to false positives, combining approaches is an effective strategy for identifying true outlier loci (de Villemereuil et al. 2014; François et al. 2016; Liggins et al. 2019). Here, we retained only loci identified as outliers by at least three methods and divided each of our datasets into neutral and outlier loci.

Analysis of Neutral and Adaptive Population Structure

We estimated several metrics of genetic diversity for each population (percentage of polymorphic loci, frequency of the most common [major] allele, expected (H_E) and observed (H_O) heterozygosity) as well as F_{IS} for each neutral and outlier dataset using the *populations* program in STACKS. We also estimated pairwise population differentiation (F_{ST}) for each dataset and assessed the significance of these estimates using permutations (999) with GenoDive (v3.04; Meirmans 2020).

We assessed population structure of neutral and outlier datasets using two approaches in R. First, we conducted a discriminant analysis of principal components (DAPC) using the package adegenet (v2.1.3; Jombart et al. 2010), which creates synthetic axes that maximize between-K and minimize within-K variance. The optimal K was chosen using the Bayesian Information Criterion (BIC), and the alpha-score was used to determine the number of PCs to retain so as not to incur overfitting issues. We also used the spatially explicit method TESS3, implemented in the package tess3r (v1.1.0; Caye et al. 2016), which estimates global ancestry coefficients while considering spatial proximity among individuals. For TESS3, we chose the optimal K using the cross-validation score.

To test for the potential influence of isolation by distance (IBD) we plotted the relationship between the shortest marine distance and pairwise neutral F_{ST} and assessed the significance of their linear relationship using the *cor.test* function of the stats package in R. We also performed Mantel tests between pairwise neutral F_{ST} and shortest marine distance using the package vegan (v2.5.6; Oksanen et al. 2019) and evaluated significance with 999 permutations. As our data were non-normally distributed (Shapiro-Wilk normality tests $p<0.05$; Shapiro & Wilk 1965), we used the non-parametric Spearman's r statistic for all tests. Geographic distances were represented as the shortest marine distance between each pair of sites, estimated using least cost path analysis in ArcGIS, for which land cells were categorized as impermeable barriers.

Hydrodynamic Model used for Larval Dispersal Modelling

We used hydrodynamic models to estimate expected larval connectivity in*S. umbilicalis* . Simulated three-dimensional ocean currents from these models were used to advect passive particles in the Lagrangian larval dispersal model described below. The hydrodynamic model outputs were from the NEMO Atlantic Margin Model (AMM15; Graham et al. 2018), which has a model horizontal resolution of 1.5 km and 51 terrain-following vertical layers. Simulated flow fields are available at hourly intervals (see Graham et al. 2018 for details of evaluation) and the flow fields from June-Sept 2014 were used to encompass the spawning season of *S. umbilicalis* . Coscia et al. (2020) demonstrated little interannual variability in the patterns of connectivity matrices of larvae advected using AMM15 output (2008-2014), hence, the background connectivity for 2014 was taken as the annual connectivity. This annual background connectivity between the spawning sites (described below) was used for calculating multi-year spread.

Particle-tracking Estimates of Larval Connectivity

Virtual particles representing larvae were released from 438 distinct release sites distributed equidistantly along the Irish Sea coastline, at 10 km intervals and 1 km offshore (Fig. 1B). By selecting release sites 1 km offshore, all particles were released within the hydrodynamic model grid cell adjacent to the coast, appropriate for simulating spawning of near-shore species. Within a 100 m radius of each release site, 625 release locations were randomly spaced to allow for some stochasticity within the simulated currents and resulting particle advection that represents unresolved sub-grid-scale and sub-time-scale processes. Cohorts of 625 particles were released daily (at 12:00 noon) from each of the 438 sites. To encompass releases during the full spring-neap tidal cycle, particles were released for the first 16 days of June. To capture dispersal variability over seasonal timescales, this procedure was repeated for July, August and September, which covers the spawning season of *S. umbilicalis* (Williams 1964; Underwood 1972; Garwood & Kendall 1985). Adopting a similar experimental design to this study, Robins et al. (2013) showed that release of 10,000 particles was sufficient to represent dispersal and connectivity patterns within the Irish Sea. Here, 40,000 individual particles were released from each site, totalling 17.52 million particles (625 particles $\times$ 438 sites $\times$ 16 days $\times$ 4 months) within the Irish Sea. Particles were allowed to propagate for 10 days from release, based on the maximum pelagic larval duration of *S. umbilicalis* (Keith et al. 2011). During each 10-day simulation, individual particles were advected, and their locations were recorded hourly.

Given that larvae have been shown to settle passively in areas of low current velocity (Kendall & Lewis 1986), the particles were parameterized as passive, neutrally buoyant, and released at the surface to simulate maximum dispersal. No criteria were applied to account for habitat type, and larvae were able to settle at any of the 438 sites. Sexual maturity in *S. umbilicalis* can occur between 1-2 years of age (Williams 1964; Bode et al. 1986; Garwood & Kendall 1985), but for model simplicity we assumed sexual maturity occurred within 1 year. We did not consider mortality/fecundity of larvae so that our results represent the maximum dispersal potential.

Pairwise site connectivity for the 12 sampled sites was determined by calculating the proportion of particles released from a source site that arrived within a 5 km radius of a settlement site. This was calculated at each model time-step, between 2-10 days after their release, producing an overall averaged connectivity. The 5 km radius was set to minimize overlap between release and settlement areas, but also to account for dispersal potential over a tidal cycle based on the average tidal excursion of the Irish Sea (Robins et al. 2013).

Multi-generational connectivity was estimated between all 438 sites by accounting for the connectivity of each site with all other sites to quantify the year-on-year spread, over 100 years (Fig. S1). In this way, the intermediary sites implemented between our 12 sampled sites facilitated stepping-stone spread along coastal sites of the Irish Sea. Here, we have focused on potential larval spread via stepping-stone connectivity and have thus assumed the metapopulation size remains constant, with population fecundity equalling mortality (also see Giménez et al. 2020b). Pairwise connectivity estimates were averaged over the four months that the model was run to create a connectivity matrix between each pair of sites, and this year-on-year spread was calculated for 100 years. The matrix used in subsequent analyses was selected based on the maximum number of years it took for any pair of sites to become connected, excluding sites that never became connected over the 100-year timeframe (Fig. S2, Table S3).

Estimating Spatial & Dispersal Distances

Pairwise geographic and larval dispersal distances were converted into site-based metrics to be used as independent variables in our redundancy analysis (RDA). To represent geographic structure, we estimated distance-based Moran's eigenvector maps (dbMEMs; Dray et al. 2006). Here, site latitude and longitude were transformed into Cartesian coordinates using the *geoXY* function of the R package SoDA (v1.0.6.1; Chambers 2020), which were used to calculate a matrix of Euclidian distances using the *dist* function of the stats package. We then used the *pcnm* function of the package vegan to transform the spatial distances into rectangular matrices that describe spatial structure at multiple scales suitable for constrained ordination analyses (Borcard & Legendre 2002; Peres Neto & Legendre 2010).

To account for the asymmetric directionality of ocean current-mediated larval dispersal between sites we translated larval connectivity matrices produced by the larval dispersal modelling into asymmetric eigenvector maps (AEMs; Blanchet et al. 2008, 2011). Connectivity matrices were used to construct a node-to-edge matrix representing direct and indirect connectivity between our 12 sites using the*aem.build.binary* function of the adespatial package in R (v0.3.8; Dray et al. 2020) (i.e., an edge is present [1] if the probability of larval connectivity between sites is >0, and not present [0] if the probability of larval connectivity is 0). This resulted in a node-to-edge matrix comprised of our 12 sites (nodes) and 66 edges (connectivity links between sites). This pairwise binary matrix was translated into site-based AEM vectors using the AEM package (v0.6; Blanchet et al. 2015), with weights applied to each edge representing the average probability of connectivity from the larval connectivity matrix, which was transformed using min-max normalization.

Multivariate redundancy analysis (RDA): assessing spatial & environmental influences on genetic structure

We used RDAs to estimate the relative influence of spatial and environmental factors on neutral and putatively adaptive genetic structure in *S. umbilicalis* and *N. lapillus* . Our response variables were minor allele frequencies (MAF) of each locus, estimated using the software PLINK (v1.9; Chang et al. 2015), and detrended using the Hellinger method implemented in the *decostand* function of vegan. To account for our large number of molecular markers, we conducted PCAs on each neutral and outlier dataset and retained only meaningful PCs (those with eigenvalues >1) as response variables in our models. Environmental (PCs) and spatial (dbMEM and AEM vectors) variables were used as predictor variables. We tested for correlations between these predictors and removed one variable when correlation exceeded 0.7, resulting in a final predictor variable dataset of four environmental PCs representing SST, AT and exposure, six dbMEMs, and three AEMs. As *N. lapillus* is a direct developer, we excluded the AEM vectors from our RDA models for this species.

We conducted RDA and partial RDA analyses on our four response variable datasets (neutral and outlier datasets for *N. lapillus* and*S. umbilicalis*) using vegan. First, we conducted a backwards and forwards selection procedure using the *ordistep* function to determine the combination of predictor variables that best explained each of our response variable datasets (i.e., the model producing the highest adjusted R^2). From this "best" model, we conducted partial RDAs, where we conditioned the model to control for the influence of either geographic structure (dbMEMs), larval connectivity (AEMs) or environmental variation (PCs) by first estimating and removing their effects and then performing an RDA on the residual matrix.

Thereby, we were able to partition the variance of our "best" models, to determine the amount of explainable variation in our dataset attributed to each set of predictor variables, while controlling for all other variables. The significance of our models and associated predictor variables were tested using analyses of variance (ANOVAs), implemented using the *anova* function of the stats package in R, with 1,000 permutations.

Outlier Gene Annotation

We obtained consensus sequences for each outlier locus using the*populations* program of STACKS. Consensus sequences were compared to NCBI's GenBank nucleotide database (NCBI Resource Coordinators 2018), and the *N. lapillus* transcriptome (Chu et al. 2014) using the BLASTN 2.12.0+ (Camacho et al. 2009) function to identify closely matching sequences that may provide information on locus identity. Matches to NCBI's GenBank nucleotide database were filtered to include only those with percent identities between 70-100% and E-values between 0 and 1E-06, and the top matching sequence (based on E-value) for each locus was recorded. From these results, we identified loci that were annotated to coding regions. Matches to the *N. lapillus*transcriptome were annotated by comparing transcriptome contig sequences that matched an outlier locus to the UniProtKB Swiss-Prot database (The Uniprot Cosortium 2021) using BLASTN, with percent identity and E-value cutoffs of 50% and 1E-06, respectively.

RESULTS

Genomic Datasets

Our final filtered datasets comprised 11,600 and 3,844 polymorphic SNPs in *S. umbilicalis* and *N. lapillus* , respectively (Table 1). The filtered *S. umbilicalis* dataset contained 196 individuals, with an average of 16.33 samples per site (range of 6-19), and an average of 11.4% missing data across all populations (range of 2.1-29.2%; Table S4). The filtered *N. lapillus* dataset contained 200 samples, with an average of 16.67 samples per population (range of 13-19), and an average of 11.7% missing data across all populations (range of 1.8-29.9%; Table S4).

There were significantly fewer polymorphic loci in the *N. lapillus* dataset, which largely resulted from the filtering pipeline removing 93.35% of SNPs with mean read depths <10 (in comparison to 57.66% in the *S. umbilicalis* dataset). This result highlights an important limitation of RADseq datasets, in that they only capture a portion of genome-wide variation (Lowry et al. 2017a). Consequently, species with larger genomes will have fewer genomic regions sequenced at sufficient coverage for reliable variant detection. This suggests that *N. lapillus* has a much larger genome size than*S. umbilicalis* . Further, as gene content does not scale proportionally with genome size (Lowry et al. 2017a,b), the presumed larger genome of *N. lapillus* should also result in fewer coding sequences and the detection of fewer adaptive SNPs, which was also evident here.

Environmental & Spatial Datasets

We retained four environmental PCs, explaining 92.6% of variance in the raw environmental data (Table 2; see also Table S5 for estimations of variance, eigenvalues, and loadings). The first PC described winter AT (mean winter AT and minimum AT of the coldest month). PC2 was best described by summer AT and SST (mean and minimum summer SST, maximum summer AT, and maximum AT of the warmest month). The third PC was largely representative of wave exposure, but also included annual temperature range. PC4 was described by a combination of summer AT and SST (maximum summer SST, maximum AT of the warmest month), as well as annual temperature range.

Estimations of spatial (dbMEM) and dispersal (AEM) distances produced eight dbMEM vectors representing nonlinear gradients in spatial structure (Table 2, Fig. S3), and three AEM vectors representing current-mediated larval connectivity (Table 2).

Outlier Detection

For *S. umbilicalis* , 143 loci (1.2%) were identified by at least three methods, with 52 and 14 of these loci identified by four and five methods, respectively (Table 3). Observed heterozygosity was lower than expected and F_{IS} was positive for all populations (Table S6). F_{ST} values ranged from 0-0.368 at outlier loci compared to

0-0.026 at neutral loci, and 80.3% and 81.8% of these pairwise comparisons were significant (p=0.001) in the neutral and outlier datasets, respectively (Fig. 2A; Table S7). For *N. lapillus* , 24 loci (0.6%) were identified by at least three methods, with only one locus identified by four methods (Table 3). Observed heterozygosity was higher in *N. lapillus* and values were similar to expected heterozygosity across all sites (Table S8). F_{IS} was lower in *N. lapillus* in comparison to*S. umbilicalis* (Table S8). F_{ST} values ranged from 0.015-0.637 in outlier loci compared to 0.033-0.362 in neutral loci, and all pairwise comparisons were significant (p=0.001), with the exception of the north Wales sites (Llanddulas & Great Orme West, p=0.211) in the neutral marker dataset (Fig. 3A, Table S9). All downstream analyses of adaptive population structure were carried out on these outlier datasets. Our resulting neutral marker datasets contained 11,457 and 3,820 loci for *S. umbilicalis* and *N. lapillus* , respectively.

Neutral and Adaptive Population Structure

In the *S. umbilicalis* neutral marker dataset neither DAPC nor TESS3 were able to discern any genetic clusters using the cross-validation score, supporting K=1. Although several estimates of pairwise F_{ST} were significantly greater than zero (Fig. 2A), the uniformly low values across all population comparisons (Table S7) support the DAPC/TESS3 results indicating no substantial genetic structure. For the outlier dataset, both methods identified four as the optimal number of clusters (Fig. 2B-C). At K=4 both methods largely agreed on the geographic distribution of clusters: 1) sites in Wales, 2) Scotland and north-west England, 3) Irish Sea coasts of Ireland and Northern Ireland, and 4) south and south-eastern Ireland, with DAPC suggesting larger proportions of admixture in Irish populations. These results agree with estimates of F_{ST} (Fig. 2A; Table S7), suggesting that the largest degree of differentiation in outlier loci is between the Welsh populations and all other populations.

In the *N. lapillus* neutral marker dataset, DAPC identified five as the optimal number of genetic clusters (Fig. 3D), whereas TESS3 suggested 11 (Fig. 3B). At K=5, the genetic groups largely corresponded to the spatial distribution of sites comprising: 1) south-western Scotland, 2) Northern Ireland, 3) south and south-eastern Ireland, 4) north-western England and north Wales, and 5) west Wales and eastern Ireland (Fig. 3C). The exception of these spatial groupings is Barry, a south Wales site, which clustered with sites in north Wales and north-western England. DAPC identified limited unidirectional admixture from the Welsh/English cluster to the Irish Sea and south-eastern Ireland. At the higher level of structure, K=11, most clusters corresponded to a single site, except for two clusters comprising both Scottish sites and both sites in north Wales (Fig. 3B). In the outlier dataset, TESS3 and DAPC suggested eight and seven genetic clusters, respectively, and largely agreed on the distribution of these genetic clusters (Fig. 3C, E). Estimates of F_{ST} suggest that Welsh populations are generally the most differentiated in both neutral and outlier datasets (Fig. 3A; Table S9).

Although we detected a significant linear relationship between shortest marine distance and neutral F_{ST} in both species (Fig. 4), Mantel tests suggested a marginally significant effect of IBD in the*S. umbilicalis* dataset only (*N. lapillus* : Spearman's r=0.1485, p=0.153; *S. umbilicalis* : Spearman's r=0.2346, p=0.048).

Seascape Genomics

Neutral population structure in *N. lapillus* (N=3,820 loci) was summarized into 11 PCs that explained 100% of the variation in the dataset. The best model describing neutral population structure included two environmental predictor variables (env_PC1: p=0.0104; env_PC2: p=0.0129, Table 2). This model was significant, but only explained a small proportion of variation in the genetic dataset (p=0.0005, R^2adj=0.1862; Table 4; Fig. 5). The outlier dataset (N=24 loci) was summarized into five PCs, explaining 87.35% of the variation in the dataset. Here, the best model contained three environmental predictor variables (env_-PC1: p=0.0100; env_PC2: p=0.0003; env_PC3: p=0.0019; Table 2) and one geographic spatial variable (dbMEM_2: p=0.0016; Table 2; Fig. 6). This model was also significant, but in contrast to the neutral model, explained a larger proportion of the variation in the genetic dataset (p=0.0001, R^2adj=0.5756; Table 4; Fig. 5). Partial RDAs attributed most of this variation to the environmental predictors (R^2adj=0.2975, p=0.0022), with a smaller and non-significant proportion attributed to the geographic vector (R^2adj=0.0004, p=0.4369).

Neutral population structure in *S. umbilicalis* (N=11,457 loci), was summarized into 11 PCs that explained 100% of the variation in the dataset. The best model describing neutral population structure included two environmental predictor variables (env_PC2: p=0.0091; env_PC3: p=0.0299; Table 2), two geographic spatial variables (dbMEM_2: p=0.0366; dbMEM_3: p=0.0037; Table 2; Fig. 6), and one larval connectivity vector (AEM5: p=0.0268; Table 2; Fig. 6). This model was significant, but only explained a small proportion of variation in the genetic dataset (p=0.0013, R^2adj=0.1739; Table 4; Fig. 7). The largest proportion of explainable variation was attributed to the geographic vectors (R^2adj=0.0422, p=0.2020), with smaller proportions attributed to the larval connectivity vector (R^2adj=0.0281, p=0.2874) and environmental variables (R^2adj=0.0207, p=0.2969), however, none of these partitioned models were significant. The outlier dataset (N=143 loci), was also summarized into 11 PCs, explaining 100% of the variation in the dataset. The best model contained two geographic spatial vectors (dbMEM_2: p=0.0142; dbMEM_3: p=0.0016; Table 2; Fig. 6) and one larval connectivity vector (AEM5: p=0.0356; Table 2). This model was also significant and explained a relatively large proportion of the variation in the genetic dataset (p=0.0009, R^2adj= 0.4559; Table 4; Fig. 7). Most of the explainable variation in this model was attributed to geographic vectors (R^2adj=0.3940, p=0.0008), with a much smaller and non-significant proportion attributed to the larval dispersal vector (R^2adj=0.0459, p=0.1353).

Outlier Gene Annotation

For the 24 outlier loci identified in the *N. lapillus* dataset, four (16.7%) matched sequences in the NCBI GenBank nucleotide database using our search criteria (Table S10). Of these, two matched to the coding sequence of the *N. lapillus* estrogen receptor gene (Table 5). An additional seven outlier loci matched the *N. lapillus*transcriptome, of which, only one matched the UniprotKB Swiss-Prot database with >50% identity, an uncharacterized protein from the Golden apple snail (*Pomacea canaliculate* ; Table 5).

For the 143 outlier loci identified in the *S. umbilicalis*dataset, 11 (7.7%) matched sequences within the NCBI GenBank database using our search criteria (Table S10. Of these 11 loci, four matched sequences contained within coding regions: two within the predicted islet cell autoantigen one like gene (*ICA1L*) of the African grass rat (*Arvicanthis niloticus*), one within the estrogen receptor gene of the dogwhelk (*N. lapillus*), and one within the predicted cilia- and flagella-associated protein 100 gene (*CFAP100*) of the California sea hare (*Aplysia californica;* Table 5). We were not able to annotate any further outlier loci using the *N. lapillus*transcriptome.

DISCUSSION

We implemented a seascape genomics approach to test for differences in neutral and adaptive genetic connectivity between two intertidal marine gastropods, the direct developer *Nucella lapillus* , and the broadcast spawning *Steromphala umbilicali* s, within similarly distributed ranges in the UK and Ireland. Neutral genetic structure conformed to expectations of lower connectivity in the direct developing species, with pairwise F_{ST} being on average 11.4x higher across *N. lapillus* populations than equivalent *S. umbilicalis* populations. These findings agree with other studies comparing direct and indirect development as reproductive strategies across marine species (e.g., Sherman et al. 2008; Hoffman et al. 2011b), although reports of contrasting findings also exist (Ayre & Hughes 2000; Richards 2007). Putative outlier locus datasets identified more extensive genetic structure in both species than was identified by neutral genetic loci, suggesting a role for adaptive divergence despite ongoing gene flow. These results contribute to a growing body of literature showing that high gene flow and adaptive divergence can co-occur in marine organisms (e.g., Gleason & Burton 2016; Hoey & Pinsky 2018; Sandoval-Castillo et al. 2018; Xuereb et al. 2018; Selmoni et al. 2020). Seascape genomic modelling indicated a greater role of environment on genetic structure in *N. lapillus* than *S. umbilicalis* , with the latter being more attributable to spatial geographic patterns that coincide with the seasonal occurrence of oceanic fronts in the Irish Sea.

As hypothesized, patterns of genetic structure indicated greater connectivity across our study area in the broadcast spawning *S. umbilicalis* than the direct developer *N. lapillus* , regardless of the size of habitat gaps between sites. Although there were significant but low levels of differentiation (F_{ST}) between many sites

and a marginally significant effect of IBD, there was a lack of distinct neutral genetic structure across the sampled area for *S. umbilicalis* . Larval connectivity (AEM vectors), geographic structure (dbMEM vectors), and environmental variables (PCs representing summer AT and SST, and wave exposure) best explained the variation in this dataset, with the largest proportion of explainable variation attributed to geographic structure. However, when partitioned, no partial RDA models were significant, supporting a general lack of explainable structure in the dataset. Taken together, the genetic homogeneity of*S. umbilicalis* populations within our study area can at least partially be attributed to a combination of predicted larval dispersal over small and medium distances and multi-generational stepping-stone dispersal over larger distances. Surprisingly, predicted larval connectivity accounted for only a small proportion of variation in the neutral genetic dataset, perhaps because we did not account for interannual variability or climate change in larval connectivity matrices. Although Coscia et al. (2020) show interannual variability to be low over short time-scales (6 years), incorporating this variability may become more important over longer time-scales like those considered here (100 years). Further, what we considered to be large habitat gaps in our sampling design seemingly only constitute medium-scale dispersal distances for *S. umbilicalis* , given that genetic connectivity across these gaps was high.

Although we found much more genetic structure in *N. lapillus* than*S. umbilicalis* , we found greater connectivity between populations separated by large habitat gaps than would be expected for a direct developing species (e.g., sites 2/3 and 10; Fig. 3). This finding supports the contribution of drifting/rafting (e.g., Colson & Hughes 2004), and stepping-stone dispersal (e.g., Crandall et al. 2012) to population connectivity in this species. Neutral genetic structure in*N. lapillus* indicated 5-11 genetic clusters of geographically proximate sites, and while there was a significant relationship between shortest marine coastal distance and F_{ST}, no significant effect of IBD was detected. Neutral genetic structure was best explained by two environmental PCs representing winter AT and summer AT and SST. The effects of environmental gradients (e.g., SST, AT) on putatively neutral SNP loci may be explained by either isolation by adaptation (IBA; Nosil et al. 2009), whereby strong ecological selection against immigrants results in adaptive population divergence that restricts gene flow and allows neutral loci to diverge via genetic drift (Thibert-Plante & Hendry 2010), or loose linkage to loci under selection (Gagnaire et al. 2015). Additionally, total explainable variation was low for this dataset, supporting the contribution of additional factors to describing neutral genetic structure in this species. Future efforts would benefit from incorporating simulations of putative long-distance dispersal of *N. lapillus* (e.g., rafting).

Interestingly, sites within habitat gaps were more likely to be colonized by *N. lapillus* , which we often found in small patches or sub-optimal habitat that did not support populations of *S. umbilicalis* . This may suggest that *N. lapillus* is more of a habitat generalist than *S. umbilicalis* , or, alternatively, that direct developers will have a greater likelihood of establishing new populations since founders can be a fertilized female or a drifting egg mass, allowing multiple offspring to hatch within the same area (Johannesson 1988). This tactic facilitates rapid population increase as encounter rates between individuals will be high in species with low mobility and high rates of self-recruitment. In contrast, while *S. umbilicalis* was found at two sites within the north Wales–Scotland gap, we found only one individual at one of these sites (St. Bees Head), and a very small population in the other (Selker Bay). This suggests that while larval dispersing species may reach distant sites more quickly and frequently than direct developers, they are likely to be spread over a much broader area during their planktonic larval feeding stage (Johannesson 1988). This dispersal tactic can lead to a low density of individuals at sites beyond a critical distance, and thus low population sizes and reduced encounter rates for subsequent generations of reproduction.

In contrast to a lack of neutral genetic structure in *S. umbilicalis* , outlier loci distinguished four genetic clusters indicating that selection can still be a major driver of spatial genomic structure, even in the face of extensive gene flow. These results are not surprising for a broadcast spawning species, for which large populations sizes are expected to reduce the effects of genetic drift and thus increase the probability that population differentiation results from localized natural selection (Nielsen et al. 2009; Gagnaire et al. 2015). The distribution of adaptive genetic structure in *S. umbilicalis* corresponds to clusters separated by the habitat gaps we investigated. This structure was significantly explained by both geographic and larval dispersal

variables, but none of the environmental variables investigated. However, larval dispersal vectors were not significant in partial RDA models. Rather, most of the variation was attributed to geographic vectors representing large-scale spatial patterns differentiating sites in the northern Irish Sea/North Channel from all others (Fig. 6, dbMEM_2), and Irish from British (particularly south Wales) sites (Fig. 6, dbMEM_3). Interestingly, patterns in the spatial predictor dbMEM_2 closely resemble the "Forbes Line," which demarcates the general northern limit of southern species within the UK and Ireland (Forbes 1858), and in the summer coincides with tidal fronts in the Irish Sea (Simpson & Hunter 1974). Where they occur, fronts have been shown to affect the dispersal (e.g., Ayata et al. 2010; Firth et al. 2021), and survival (e.g., Gaylord & Gaines 2000) of species by establishing spatially stratified environmental conditions in temperature and/or salinity (Pingree et al. 1974; Pineda 1994). The extent to which these fronts may drive adaptation in *S. umbilicalis* at its northern range edge is unknown, but seasonal fronts that form in the Irish Sea during summer (Simpson et al. 2009) may be especially influential, as they coincide with the timing of spawning and dispersal. Alternatively, these results may suggest that our models did not include other important environmental variables driving selection in*S. umbilicalis* or that selective pressures may vary spatio-temporally, resulting in patterns of chaotic genetic patchiness that do not easily correlate with environmental features.

The outlier locus dataset for *N. lapillus* suggested the potential for 7-8 adaptive genetic clusters that do not correspond completely with the structure observed in the neutral dataset. This putative adaptive structure was best described by the same environmental variables as for the neutral dataset (winter AT and summer AT and SST) but with the addition of wave exposure and a minor, non-significant contribution of one geographic vector. The significance of temperature to the adaptive structure of *N. lapillus* populations in our study area corroborates previous evidence supporting thermal-mediated selection in this species. Specifically, Chu et al. (2014) identified several fixed single nucleotide polymorphisms (SNPs) within heat stress-mediated genes between clades of *N. lapillus* in the northwestern Atlantic. Additionally, the significance of wave exposure corroborates many previous studies establishing relationships between exposure and adaptation of shell morphology in *N. lapillus* populations from Europe (Hughes & Taylor 1997; Guerra-Varela et al. 2009; Pascoal et al. 2012; Carro et al. 2019) and North America (Etter 1988, 1996).

The most supported driver of environmental adaptation in marine species is temperature (Liggins et al. 2019), and in the present study we provide further evidence to substantiate its important role. Temperature is as a strong stressor in the intertidal, where it underlies many important ecological (e.g., latitudinal distributions [Helmuth et al. 2006] and vertical zonation [Somero 2002]) and physiological (Tomanek & Helmuth 2002) processes. Intertidal species are influenced by both sea and air temperatures during high and low tides, respectively. Unfortunately, detailed and consistent meteorological records are generally unavailable for intertidal regions, where sea temperatures are likely influenced by air and ground temperature, substrate, aspect, and tidal range, among other factors. For this study we used average estimates of SST and AT over large, buffered areas (10 km), and thus acknowledge that we have likely not been able to characterize fine-scale environmental features of the local seascape that may substantially influence adaptation and fitness in our species. This may partially explain the lack of environmental associations detected in our *S. umbilicalis* outlier dataset, despite the known impacts of temperature on its dispersal and physiology.

S. umbilicalis spawns multiple times throughout the year in its range centre where sea temperatures are warmer, with shorter breeding periods towards its northern range edge (the current study area) where recruitment failure is associated with cooler temperatures (Bode et al. 1986; Kendall & Lewis 1986). Given the importance of temperature on larval dispersal and settlement in *S. umbilicalis* , fine-scale temperature data may aid in identifying environmental associations with outlier loci that are involved in reproductive processes.

Our inability to identify sequence matches for most of our outlier loci is not surprising, given the distant relationship of *N. lapillus*and *S. umbilicalis* to most model species, and the underrepresentation of genomic resources and sequencing efforts for marine invertebrates (Lopez et al. 2019). Of loci that we were able to match to NCBI sequences, few were within coding regions, an increasingly common result in SNP-mapping studies. Indeed, SNPs can be as far as two Mbp away from, and are not necessarily closest to the genes they

affect (Brodie et al. 2016). Of the few loci that were located within coding regions, both species showed matches to the estrogen receptor (ER) gene. Although the function of the ER gene in molluscs is still largely uncharacterized, it has been shown to overexpress in *N. lapillus*in the presence of estrogenic chemicals in raw urban/industrial effluent pollutants, with concomitant increases in reproductive maturation (Castro et al. 2007). The other two outlier loci that mapped to coding regions in *S. umbilicalis* are less well characterized but may also be involved in reproduction (He et al. 2015; Robay et al. 2018).

Our results support the hypothesis that substantial heterogeneity of the seascape can support connectivity of the otherwise low mobility direct developing species, *Nucella lapillus* , and create adaptive divergence in an otherwise highly connected meta-population of the broadcast spawning species, *Steromphala umbilicalis* . Ultimately, characterizing spatial patterns in connectivity, estimating genetic diversity, and identifying locally adapted populations will lead to a better understanding of the resilience of marine species to changing environments and allow for improved management of fisheries and marine protected areas (MPAs; e.g., Miller & Ayre 2008; Sinclair-Waters et al. 2018). Sea temperatures around the UK and Ireland are predicted to catch up with global trends within the next decade, and associations between species abundance and sea temperatures observed over less than a decade indicate the sensitivity of many species to these changes (Mieszkowska et al. 2020). Thus, understanding the impacts of ocean warming on intertidal marine species, and their potential to adapt to these changes is an important goal.

Acknowledgements: This research was funded by the Ecostructure project (part-funded by the European Regional Development Fund (ERDF) through the Ireland-Wales Cooperation Programme 2014-2020). The authors would like to acknowledge all members of the Ecostructure project (http://www.ecostructureproject.eu/) who helped in the sample collection and processing for this work, and Data2Bio LLC (Ames, Iowa) who provided the sequencing services for this project. We are grateful to Niall McKeown, Peter Lawrence and Ally Evans for helpful discussions of the data and results. Data analysis was supported by the Supercomputing Wales programme, which is part-funded by the European Regional Development Fund (ERDF) via the Welsh government.

Data Accessibility & Benefit Sharing

The following datasets are available in the PURE repository (DOI: 10.20391/72b642b1-fa78-4568-832d-5e024315b9ae): (1) *S. umbilicalis* neutral SNP dataset (in VCF format), (2)*S. umbilicalis* outlier SNP dataset (in VCF format), (3) *N. lapillus* neutral SNP dataset (in VCF format), (4) *N. lapillus*outlier SNP dataset (in VCF format), and (5) environmental data, dbMEM vectors and AEM vectors calculated for each site. Scripts for the seascape genomics analyses are also provided in the PURE repository.

Benefits from this research accrue from the sharing of our data and results on public databases as described above.

Author Contributions

MBP, SLW, PER, SRJ, PWS and JEI conceived the ideas and designed the methodology. MBP, JEI and HSE collected the samples. MBP and HSE conducted the laboratory work. SLW conducted the larval dispersal modelling. MBP conducted the genomic analyses. CS and IS provided bioinformatic support. MBP wrote the paper and all authors contributed to critically reviewing the paper. MBP and CS produced reproducible scripts for the analyses and archived the data.

Literature Cited

Allendorf, F. W., Hohenlohe, P. A., & Luikart, G. (2010). Genomics and the future of conservation genetics. *Nature Reviews Genetics* , 11, 697-709. doi: https://doi.org/10.1038/nrg2844

Andrews, S. (2010). FastQC: a quality control tool for high throughput sequence data. Retrieved from http://www.bioinformatics.babraham.ac.uk/projects/fastqc

Ayata, S. D., Lazure, P., & Thiébaut, É. (2010). How does the connectivity between populations mediate range limits of marine invertebrates? A case study of larval dispersal between the Bay of Bis-

cay and the English Channel (North-East Atlantic). *Progress in Oceanography* , 87, 18–36. doi: https://doi.org/10.1016/j.pocean.2010.09.022

Ayre, D. J., & Hughes, T. P. (2000). Genotypic diversity and gene flow in brooding and spawning corals along the great barrier reef, Australia.*International Journal of Organic Evolution* , 54, 1590-1605. doi: https://doi.org/10.1111/j.0014-3820.2000.tb00704.x

Baco, A. R., Etter, R. J., Riberio, P. A., von der Heyden, S., Beeril, P., & Kinlan, B. P. (2016). A synthesis of genetic connectivity in deep-sea fauna and implications for marine reserve design.*Molecular Ecology* , 25, 3276-3298. doi: https://doi.org/10.1111/mec.13689

Ballantine, W. J. (1961). A biologically-defined exposure scale for comparative description of rocky shores. *Field Studies* , 1, 1-19.

Bell, J. J., & Okamura, B. (2005). Low genetic diversity in a marine nature reserve: re-evaluating diversity criteria in reserve design.*Proceedings of the Royal Society B: Biological Sciences* , 272, 1067-1074. doi: https://doi.org/10.1098/rspb.2005.3051

Bell, J. J. (2008). Similarity in connectivity patterns for two gastropod species lacking pelagic larvae. *Marine Ecology Press Series* , 357, 185-194. doi: 10.3354/meps07301

Benestan, L., Gosselin, T., Perrier, C., Sainte-Marie, B., Rochette, R., & Bernatchez, L. (2015). RAD genotyping reveals fine-scale genetic structuring and provides powerful population assignment in a widely distributed marine species, the American lobster (Homarus americanus).*Molecular Ecology* , 24, 3299-3315. doi: 10.1111/mec.13245

Benestan, L., Quinn, B. K., Maaroufi, H., Laporte, M., Clark, F. K., Greenwood, S. J., ... Bernatchez, L. (2016). Seascape genomics provides evidence for thermal adaptation and current-mediated population structure in American lobster (*Homarus americanus*).*Molecular Ecology* , 25, 5073-5092. doi: 10.1111/mec.13811

Bernatchez, S., Xuereb, A., Laporte, M., Benestan, L., Steeves, R. Laflamme, M., ... Mallet, M. A. (2019). Seascape genomics of eastern oyster (*Crassostrea virginica*) along the Atlantic coast of Canada. *Evolutionary Applications* , 12, 587-609. doi: 10.1111/eva.12741

Blanchet, F.G., Legendre, P., & Borcard, D. (2008). Modelling directional spatial processes in ecological data. *Ecological Modelling* , 215, 325-336. doi: https://doi.org/10.1016/j.ecolm odel.2008.04.001

Blanchet, F. G., Legendre, P., Maranger, R. Monti, D., & Pepin, P. (2011). Modelling the effect of directional spatial ecological processes at different scales. *Oecologia* , 166, 357-368. doi: https://doi.org/10.1007/ s00442-010-1867-y

Blanchet, F. G., Legendre, P., & Gauthier, O. (2015). AEM: tools to construct asymmetric eigenvector maps (AEM) spatial variables. R package version 0.6/r127. https://R-Forge.R-project.org/projects/sedar/

Bode A., Lombas I., & Anadon, N. (1986) Preliminary studies on the reproduction and population dynamics of *Monodonta lineata* and*Gibbula umbilicalis* (Mollusca, Gastropoda) on the central coast of Asturias (N. Spain). *Hydrobiologia* , 38, 31-39. doi: https://doi.org/10.1007/BF00026745

Borcard D., & Legendre, P. (2002). All-scale spatial analysis of ecological data by means of principal coordinates of neighbour matrices.*Ecological Modelling* , 153,51–68. doi: https://doi.org/10.1016/S0304-3800(01)00501-4

Brodie, A., Azaria, J. R., & Ofran, Y. (2016). How far from the SNP may the causative genes be? *Nucleic Acids Research* , 44, 6046-6054.

Burrows, M. T., Harvey, R., & Robb, L. (2008). Wave exposure indices from digital coastlines and the prediction of rocky shore community structure. *Marine Ecology Progress Series* , 353, 1-12. doi: https://doi.org/10.3354/meps07284

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., & Madden, T. L. (2009). BLAST+: architecture and applications. *BMC Bioinformatics* , 10, 421. doi: https://doi.org/10.1186/1471-2105-10-421

Carro, B., Quintela, M., Ruiz, J. M., & Barreiro, R. (2019). Wave exposure as a driver of isolation by environment in the marine gastropod*Nucella lapillus* . *Hydrobiologia* , 839, 51-69. doi: https://doi.org/10.1007/s10750-019-03993-5

Castle, S. L. & Emery, A.E.H. (1981). *Nucella lapillus* : a possible model for the study of genetic variation in natural populations. *Genetica* , 56, 11 - 15. doi: https://doi.org/10.1007/BF00126924

Castro, L. F. C., Melo, C., Guillot, R., Mendes, I., Queirós, S., Lima, D., ... Santos, M. M. (2007). The estrogen receptor of the gastropod *Nucella lapillus* : modulation following exposure to an estrogen effluent? *Aquatic Toxicology* , 84, 465-468.

Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A., & Cresko, W. A. (2013). Stacks: an analysis tool set for population genomics.*Molecular Ecology* , 22, 3124-3140. doi: https://doi.org/10.1111/mec.12354

Caye, K., Deist, T. M., Martins, H., Michel, O., & François, O. (2016). TESS3: fast inference of spatial population structure and genome scans for selection. *Molecular Ecology Resources* , 16, 540-548. doi: https://doi.org/10.1111/1755-0998.12471

Chambers, J. M. (2020). SoDA: Functions and examples for "software data analysis." R package version 1.0-6.1, https://CRAN.R-project.org/package=SoDA

Chang, C. C., Chow, C. C., Tellier, L. C. A. M., Vattikuti, S., Purcell, S. M., & Lee, J. J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience* , 4, 7. doi: https://doi.org/10.1186/s13742-015-0047-8

Chu, N. D., Kaluziak, S. T., Trussell, G. C., & Vollmer, S. V. (2014). Phylogenomic analyses reveal latitudinal population structure and polymorphisms in heat stress genes in the North Atlantic snail*Nucella lapillus* . *Molecular Ecology* , 23, 1863-1873.

Colson, I., & Hughes, R. N. (2004). Rapid recovery of genetic diversity of dogwhelk (*Nucella lapullus* L.) populations after local extinction and recolonization contradicts predictions from life-history characteristics. *Molecular Ecology* , 13, 2223-2233. doi: https://doi.org/10.1111/j.1365-294X.2004.02245.x

Colson, I., & Hughes, R. N. (2007) Contrasted patterns of genetic variation in the dogwhelk *Nucella lapillus* along two putative post-glacial expansion routes. *Marine Ecology Press Series* , 343, 183-191. doi: https://doi.org/10.3354/meps06879

Coscia, I., Wilmes, S. B., Ironside, J. E., Goward-Brown, A., O'Dea, E., Mlaham, S. K., ... Robins, P. E. (2020). Fine-scale seascape genomics of an exploited marine species, the common cockle*Cerastoderma edule* , using a multimodelling approach.*Evolutionary Applications* , 13, 1854-1867. doi: https://doi.org/10.1111/eva.12932

Crandall, E. D., Treml, E. A., & Barber, P. H. (2012). Coalescent and biophysical models of stepping-stone gene flow in neritid snails.*Molecular Ecology* , 21, 5579-5598. doi: https://doi.org/10.1111/mec.12031

D'Aloia, C. C., Bogdanowicz, S. M., Francis, R. K., Majoris, J. E., Harrison, R. G. & Buston, P. M. (2015). Patterns, causes, and consequences of marine larval dispersal. *Proceedings of the National Academy of Sciences of the United States of America* , 112, 13940-13945. doi: https://doi.org/10.1073/pnas.1513754112

de Villemereuil, P., Frichot, E., Bazin, E., François, O., & Gaggiotti, O. E. (2014). Genome scan methods against more complex models: when and how much should we trust them? *Molecular Ecology* , 23, 2006-2019. doi: 10.1111/mec.12705

de Villemereuil, P., & Gaggiotti, O. E. (2015). A new F_{ST}-based method to uncover local adaptation using

environmental variables. *Methods in Ecology and Evolution* , 6, 1248-1258. doi: https://doi.org/10.1111/2041-210X.12418

DiBattista, J. D., Travers, M. J., Moore, G. I., Evans, R. D., Newman, S. J. Feng, M., ... Berry, O. (2017). Seascape genomics reveals file-scale patterns of dispersal for a reef fish along the ecologically divergent coast of Northwestern Australia. *Molecular Ecology* , 26, 6206-6223. doi: 10.1111/mec.14352

Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., ... Lautenbach, S. (2013) Colinearity: a review of methods to deal with it and a simulation study evaluating their performance.*Ecography* , 36, 27-46. doi: 10.1111/j.1600-0587.2012.07348.x

Dray, S., Legendre, P., & Peres-Neto, P. R. (2006). Spatial modelling: A comprehensive framework for principal coordinate analysis of neighbour matrices (PCNM). *Ecological Modelling* , 196, 483–493. doi: https://doi.org/10.1016/j.ecolmodel.2006.02.015

Dray, S., Bauman, D., Blanchet, G., Borcard, D., Clappe, S., Guenard, G... Wagner, H. (2020) adespatial: multivariate multiscale spatial analysis. R package version 0.3-8. https://CRAN.R-project.org/package=adespatial

Edmands, S., & Potts, D. C. (1997). Population genetic structure in brooding sea anemones (Epiactis spp.) with contrasting reproductive modes. *Marine Biology* , 127, 485-498. doi: https://doi.org/10.1007/s002270050037

ESRI. (2011). ArcGIS Desktop: Release 10. Redlands, CA: Environmental Systems Research Institute.

Etter, R. J. (1988). Physiological stress and color polymorphism in the intertidal snail *Nucella lapillus* . *Evolution* , 42, 660-680.

Etter, R. J. (1996). The effect of wave action, prey type, and foraging time on growth of the predatory snail *Nucella lapillus* (L.).*Journal of the Marine Biological Association of the United Kingdom* , 196, 341–356. doi: https://doi.org/10.1016/0022-0981(95)00139-5

Ewels, P., Magnusson, M., Lundin, S., & Käller, M. (2016). MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* , 32, 3047-3048. doi: https://doi.org/10.1093/bioinformatics/btw354

Excoffier, L., & Lischer, H.E. L. (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources* , 10, 564-567. doi: https://doi.org/10.1111/j.1755-0998.2010.02847.x

Fick, S. E., & Hijmans, R. J. (2017). WorldClim 2: new 1km spatial resolution climate surfaces for global land areas. *International Journal of Climatology* , 37, 4302-4315. doi: https://doi.org/10.1002/joc.5086

Firth, L. B., Harris, D., Blaze, J. A., Marzloff, M. P., Boyé, A., Miller, P. I. ..., & Hawkins, S. J. (2021). Specific niche requirements underpin multidecadal range edge stability, but may introduce barriers for climate change adaptation. *Diversity and Distributions* , 27, 668-683. doi: 10.1111/ddi.13224

Flanagan, S. P., Forester, B. R., Latch, E. K., Aitken, S. N., & Hoban, S. (2018). Guidelines for planning genomic assessment and monitoring of locally adaptive variation to inform species conservation.*Evolutionary Applications* , 11, 1035-1052. doi: 10.1111/eva.12569

François, O., Martins, H., Caye, K., Schoville, S. D. (2016). Controlling false discoveries in genome scans for selection.*Molecular Ecology* , 25, 454-469. doi: https://doi.org/10.1111/mec.13513

Frichot, E., & François, O. (2015). LEA: an R package for landscape and ecological association studies. *Methods in Ecology and Evolution* , 6, 925-929. doi: 10.1111/2041-210X.12382

Foll, M., & Gaggiotti, O. (2008). A genome-scan method to identify selected loci appropriate for

both dominant and codominant markers: A Bayesian perspective. *Genetics* , 180, 977-993. doi: https://doi.org/10.1534/genetics.108.092221

Forbes, E. (1858). The distribution of marine life, illustrated chiefly by fishes and molluscs and radiata. In A. K. Johnston's Physical Atlas. (Ed.), (pp. 99–101). W. & A. K. Johnston.

Gagnaire, P-A., Borquet, T., Aurelle, D., Viard, F., Souissi, A., Bonhomme, F., ... Bierne, N. (2015). Using neutral, selected, and hitchhiking loci to assess connectivity of marine populations in the genomic era. *Evolutionary Applications* , 8, 769-786. doi: https://doi.org/10.1111/eva.12288

Galindo, H. M., Olson, D. B. & Palumbi, S. R. (2006). Seascape genetics: a coupled oceanographic-genetic model predicts population structure of Caribbean corals. *Current Biology* , 16, 1622-1626. doi: https://doi.org/10.1016/j.cub.2006.06.052

Garwood, P. R., & Kendall, M. A. (1985). The reproductive cycles of*Monodonta lineata* and *Gibbua umbilicalis* on the coast of mid-Wales. *Journal of the Marine Biological Association of the United Kingdom* , 65, 993-1008. doi: https://doi.org/10.1017/S0025315400019470

Gaylord, B., & Gaines, S. D. (2000). Temperature or transport? Range limits in marine species mediated solely by flow. *The American Naturalist* , 155, 769–789. doi: https://doi.org/10.1086/303357

Giménez, L., Robins, P., & Jenkins, S. R. (2020a). Role of trait combinations, habitat matrix, and network topology in metapopulation recovery from regional extinction. *Limnology and Oceanography* , 65, 775-789. doi: https://doi.org/10.1002/lno.11347

Giménez, L., Exton, M., Spitzner, F., Meth, R., Ecker, U., Jungblut, S., ..., & Torres, G. (2020b). Exploring larval phenology as predictor for range expansion in an invasive species. *Ecography* , 43, 1423-1434. doi: https://doi.org/10.1111/ecog.04725

Gleason, L. U., & Burton, R. S. (2016). Genomic evidence for ecological divergence against a background of population homogeneity in the marine snail *Chlorostoma funebralis* . *Molecular Ecology* , 25, 3557-3573. doi: https://doi.org/10.1111/mec.13703

Graham, J. A., O'Dea, E., Holt, J., Polton, J., Hewitt, H. T., Furner, R., ..., & Mayorga Adame, C. G. (2018). AMM15: a new high-resolution NEMO configuration for operational simulation of the European north-west shelf. *Geoscientific Model Development* , 11, 681-696. doi: https://doi.org/10.5194/gmd-11-681-2018

Guerra-Varela, J., Colson, I., Backeljau, T., Breugelmans, K., Hughes, R. N., & Rolán-Alvarez, E. (2009). The evolutionary mechanism maintaining shell shape and molecular differentiation between two ecotypes of the dogwhelk *Nucella lapillus* . *Evolutionary Ecology* , 23, 261-280. doi: 10.1007/s10682-007-9221-5

He, J., Xia, M., Tsang, W. H., Chow, K. L., & Xia, J. (2015). ICA1L forms BAR-domain complexes with PICK1 and is crucial for acrosome formation in spermiogenesis. *Journal of Cell Science* , 128, 3822-3836.

Helmuth, B., Mieszkowska, N., Moore, P., & Hawkins, S. J. (2006) Living on the edge of two changing worlds: forecasting the responses of rocky intertidal ecosystems to climate change. *Annual Review of Ecology, Evolution, and Systematics* , 37, 373–404. doi: https://doi.org/10.1146/annurev.ecolsys.37.091305.110149

Hoey, J. A., & Pinsky, M. L. (2018). Genomic signatures of environmental selection despite near-panmixia in summer flounder.*Evolutionary Applications* , 11, 1732-1747. doi: https://doi.org/10.1111/eva.12676

Hoffman, J. I., Pack, L. S., Linse, K., & Clarke, A. (2011a). Strong population genetic structure in a broadcast-spawning Antarctic marine invertebrate. *Journal of Heredity* , 102, 55-66. doi: https://doi.org/10.1093/jhered/esq094

Hoffman, J. I., Clarke, A., Linse, K., & Peck L. S. (2011b). Effects of brooding and broadcasting reproductive modes on the population genetic structure of two Antarctic gastropod molluscs. *Marine Biology* , 158, 287-296. doi: 10.1007/s00227-010-1558-6

Hughes, R. N., & Taylor, M. J. (1997). Genotype-environment interaction expressed in the foraging behaviours of dogwhelks, *Nucella lapillus* (L), under simulated environmental hazard. *Proceedings of the Royal Society B: Biological Sciences* , 264, 417-422. doi: https://doi.org/10.1098/rspb.1997.0059

Johannesson, K. (1988). The paradox of Rockall: why is a brooding gastropod (*Littorina saxatilis*) more widespread than one having a planktonic larval dispersal stage (*L. littorea*)? *Marine Biology* , 99, 507-513. doi: https://doi.org/10.1007/BF00392558

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. doi: https://doi.org/10.1186/1471-2156-11-94

Keith, S. A., Herbert, R. J. H., Norton, P. A., Hawkins, S. J., & Newton, A. C. (2011) Individualistic species limitations of climate-induced range expansions generated by meso-scale dispersal barriers. *Diversity and Distributions* , 17, 275-286. doi: https://doi.org/10.1111/j.1472-4642.2010.00734.x

Kendall, M. A., & Lewis, J. R. (1986) Temporal and spatial patterns in the recruitment of *Gibbula umbilicalis*. *Hydrobiologia* , 142, 15-22. doi: https://doi.org/10.1007/BF00026743

Kendall, M. A., Williamson, P., & Garwood, P. R. (1987). Annual variation in recruitment and population structure of *Monodonta lineata* and *Gibbula umbilicalis* populations at Aberaeron, mid-Wales. *Estuarine, Coastal and Shelf Science* , 24, 499-511. doi: https://doi.org/10.1016/0272-7714(87)90130-2

Kirby, R. R. (2000). An ancient transpacific polymorphism shows extreme divergence in a multitrait cline in an intertidal snail (*Nucella lapillus* (L.)). *Molecular Biology and Evolution* , 17, 1816-1825. doi: https://doi.org/10.1093/oxfordjournals.molbev.a026282

Kyle, C. J., & Boulding, E. G. (2000). Comparative population genetic structure of marine gastropods (Littornia spp.) with and without pelagic larval dispersal. *Marine Biology* , 137, 835-845. doi: https://doi.org/10.1007/s002270000412

Lenormand, T. (2002). Gene flow and the limits to natural selection. *Trends in Ecology and Evolution* , 17, 183-189. doi: https://doi.org/10.1016/S0169-5347(02)02497-7

Liggins, L., Treml, E. A., & Riginos, C. (2019). Seascape genomics: Contextualizing adaptive and neutral genomic variation in the ocean environment. In: Oleksiak, M., & Rajora, O. P. (Eds.), *Population Genomics: Marine Organisms* . (pp. 49). Switzerland: Springer Nature.

Lischer, H. E. L., & Excoffier, L. (2012). PGDSpider: an automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics* , 28, 298-299. doi: https://doi.org/10.1093/bioinformatics/btr642

Lopez, J. V., Kamel, B., Medina, M., Collins, T., & Baums, I. B. (2019). Multiple facets of marine invertebrate conservation genomics. *Annual Review of Animal Biosciences* . 7, 22.1-22.25. doi: https://doi.org/10.1146/annurev-animal-020518-115034

Lowry, D. B., Hoban, S., Kelleu, J. L., Lotterhos, K. E., Reed, L. K., Antolin, M. F., & Storfer, A. (2017a). Breaking RAD: an evaluation of the utility of restriction site-associated DNA sequencing for genome scans of adaptation. *Molecular Ecology* , 17, 142-152. doi: 10.1111/1755-0998.12635

Lowry, D. B., Hoban, S., Kelley, J. L., Lotterhos, K. E., Reed, L. K., Antolin, M. F., & Storfer, A. (2017b). Responsible RAD: striving for best practices in population genomic studies of adaptation. *Molecular Ecology* , 17, 366-369. doi: https://doi.org/10.1111/1755-0998.12677

Luu, K., Bazin, E., & Blum, M. G. B. (2017). pcadapt: an R package to perform genome scans for selection based on principal component analysis. *Molecular Ecology Resources* , 17, 67-77. doi: https://doi.org/10.1111/1755-0998.12592

Meirmans, P. G. (2020). GENODIVE version 3.0: easy-to-use software for the analysis of genetic data of diploids and polyploids. *Molecular Ecology Resources* , 20, 1126-1131. doi: 10.1111/1755-0998.13145

Mieszkowska, N., Milligan, G., Burrows, M. T., Freckleton, R., & Spencer, M. (2013). Dynamic species distribution models from categorical survey data. *Journal of Animal Ecology* , 82, 1215–1226. doi: https://doi.org/ 10.1111/1365-2656.12100

Mieszkowska, N., Burrows, M., & Sugden, H. (2020). Impacts of climate change on intertidal habitats relevant to the coastal and marine environmental around the UK. *MCCIP Science Review 2020* , 256-271. doi: 10.14465/2020.arc12.ith

Miller, K. J., & Ayre, D. J. (2008). Protection of genetic diversity and maintenance of connectivity among reef corals within marine protected areas. *Conservation Biology* , 22, 1245-1254. doi: https://doi.org/10.1111/j.1523-1739.2008.00985.x

NCBI Resource Coordinators. (2018). Database resources for the National Centre for Biotechnology Information. *Nucleic Acids Research* , 4, D8-D13. doi: 10.1093/nar/gkx1095

Nielsen, E. E., Hemmer-Hansen, J., Larsen, P. F., & Bekkevold, D. (2009). Population genomics of marine fishes: identifying adaptive variation in space and time. *Molecular Ecology* , 18, 3128-3150. doi: https://doi.org/ 10.1111/j.1365-294X.2009.04272.x

Nosil, P., Funk, D. J., & Ortiz-Barrientos, D. (2009). Divergent selection and heterogeneous genomic divergence. *Molecular Ecology* , 18, 375-402. doi: https://doi.org/10.1111/j.1365-294X.2008.03946.x

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, ... Wagner, H. (2019). Vegan: community ecology package. R package version 2.5-6. https://CRAN.R-project.org/package=vegan

Ott, A., Liu, S., Schnable, J. C., Yeh, C-T. E., Wang, K-S., & Schnable, P. S. (2017). tGBS(r) genotyping-by-sequencing enables reliable genotyping of heterozygous loci. *Nucleic Acids Research* , 45, e178. doi: 10.1093/nar/gkx853

Palumbi, S. R. (1994). Genetic divergence, reproductive isolation, and marine speciation. *Annual Review of Ecology, Evolution and Systematics* , 25, 547-572. doi: https://doi.org/10.1146/annurev.es.25.110194.002555

Palumbi, S. R. (2004). Marine reserves and ocean neighbourhoods: the spatial scale of marine populations and their management. *Annual Review of Environment and Resources* , 29, 31-68. doi: https://doi.org/10.1146/annurev.energy.29.062403.102254

Paris, J. R., Stevens, J. R., & Catchen, J. M. (2017). Lost in parameter space: a road map for STACKS. *Methods in Ecology and Evolution* , 8, 1360-1373. doi: https://doi.org/10.1111/2041-210X.12775

Pascoal, S., Carvalhom G., Creer, S., Rock, J., Kawaii, K., Mendo, S. & Hughes, R. (2012). Plastic and heritable components of phenotypic variation in *Nucella lapillus* : an assessment using reciprocal transplant and common garden experiments. *PLoS ONE* , 7:e30289. doi: https://doi.org/10.1371/journal.pone.0030289

Pascual, M., Rives, B., Schunter, C., & Macpherson, E. (2017). Impact of life history traits on gene flow: a multispecies systematic review across oceanographic barriers in the Mediterranean Sea. *PLoS One* , 12, e0176419. doi: 10.1371/journal.pone.0176419

Peres Neto, P. R., & Legendre, P. (2010). Estimating and controlling for spatial structure in the study of ecological communities. *Global Ecology and Biogeography* , 19, 174-184. doi: https://doi.org/10.1111/j.1466-8238.2009.00506.x

Pineda, J. (1994). Internal tidal bores in the nearshore: warm-water fronts, seaward gravity currents and the onshore transport of neustonic larvae. *Journal of Marine Research* , 52, 427–458. doi: https://doi.org/10.1357/00222 40943 077046

Pingree, R. D., Forster, G. R., & Morrison, G. K. (1974). Turbulent Convergent Tidal Fronts. *Journal of the Marine Biological Association of the United Kingdom* , 54, 469-479. doi: https://doi.org/10.1017/S0025315400058653

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/

Rellstab, C., Gugerli, F., Eckert, A. J., Hancock, A. M., & Holderegger, R. (2015). A practical guide to environmental association analysis in landscape genomics. *Molecular Ecology* , 24, 4348-4370. doi: 10.1111/mec.13322

Richards, V. P., Thomas, J. D., Stanhope, M. J., & Shivji, M. S. (2007). Genetic connectivity in the Florida reef system: comparative phylogeography of commensal invertebrates with contrasting reproductive strategies. *Molecular Ecology* , 16, 139-157. doi: https://doi.org/10.1111/j.1365-294X.2006.03145.x

Robay, A., Abbasi, S., Akil, A., El-Bardisi, H., Arafa, M., Crystal, R. G., & Fakhuro, K. A. (2018). A systemic review on the genetics of male infertility in the era of next-generation sequencing. *Arab Journal of Urology* , 16, 53-64. doi: 10.1016/j.aju.2017.12.003

Robins, P. E., Neill, S. P., Gimenez, L., Stuart, R., Jenkins, S. R., & Malham, S. K. (2013). Physical and biological controls on larval dispersal and connectivity in a highly energetic shelf sea.*Limnology and Oceanography* , 58, 505-524. doi: https://doi.org/10.4319/lo.2013.58.2.0505

Rolan, E., Guerra-Varela, J., Colson, I., Hughes, R. N., & Rolan-Alvarez, E. (2004). Morphological and genetic analysis of two sympatric morphs of the dogwhelk *Nucella lapillus* (Gastropoda: Muricidae) from Galicia (northwestern Spain). *Journal of Molluscan Studies* , 70, 179-185. doi: https://doi.org/10.1093/mollus/70.2.179

Sandoval-Castillo, J., Robinson, N. A., Hart, A. M., Strain, L. W. S., & Beheregaray, L. B. (2018). Seascape genomics reveals adaptive divergence in a connected and commercially important mollusc, the greenlip abalone (*Haliostis laevigata*), along a longitudinal environmental gradient. *Molecular Ecology* , 27, 1603-1620. doi: 10.1111/mec.14526.

Sanford, E., & Kelly, M. W. (2011). Local adaptation in marine invertebrates. *Annual Review of Marine Science* , 3, 509-535. doi: 10.1146/annurev-marine-120709-142756

Selkoe, K. A., Watson, J. R., Horin, T. B., Iacchei, M., Mitarai, S. Siegel, D. A., ... Toonen, R. J. (2010). Taking the chaos out of genetic patchiness: seascape genetics reveals ecological and oceanographic drivers of genetic patterns in three temperate reef species. *Molecular Ecology* , 19, 3708-3726. doi: https://doi.org/10.1111/j.1365-294X.2010.04658.x

Selkoe, K. A., & Toonen, R. J. (2011). Marine connectivity: a new look at pelagic larval duration and genetic metrics of dispersal.*Marine Ecology Press Series* , 436, 291-305. doi: https://doi.org/10.3354/meps09238

Selkoe, K. A., D'Aloia, C. C., Crandall, E. D., Iacchei, M., Liggins, L., Puritz, J. B., ... Toonen, R. J. (2016). A decade of seascape genomics: contributions to basic and applied marine connectivity.*Marine Ecology Press Series* , 554, 1-19. doi: 10.3354/meps11792

Selmoni, O., Rochat, E., Lecellier, G., Berteaux-Lecellier, V., & Joost, S. (2020). Seascape genomics as a new tool to empower coral reef conservation strategies: an example on north-western Pacific*Acropora digitifera* . *Evolutionary Applications* , 13, 1923-1938. doi: 10.1111/eva.12944

Shanks, A. L. (2009). Pelagic larval duration and dispersal distance revisited. *The Biological Bulletin* , 216, 373-385. doi: https://doi.org/10.1086/BBLv216n3p373

Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika* ,52, 591-611. doi: https://doi.org/10.1093/biomet/52.3-4.591

Sherman, C. D. H., Hunt, A., & Ayre, D. J. (2008). Is life history a barrier to dispersal? Contrasting patterns of genetic differentiation along an oceanographically complex coast. *Biological Journal of the Linnean Society* , 95, 106-116. doi: https://doi.org/10.1111/j.1095-8312.2008.01044.x

Simpson, J. H., Green, J. M., Rippeth, T. P., Osborn, T. R., & Nimmo- Smith, W. A. M. (2009). The structure of dissipation in the western Irish Sea front. Journal of Marine Systems, 77, 428–440. https://doi.org/10.1016/j.jmars ys.2008.10.014

Simpson, J. H., & Hunter, J. R. (1974). Fronts in the Irish sea.*Nature* , 250, 404-406. doi: https://doi.org/10.1038/250404a0

Sinclair-Waters, M., Bentzen, P., Morris, C. J., Ruzzante, D. E., Kent, M. P. Lein, S., & Bradbury, I. R. (2018). Genomic tools for management and conservation of Atlantic cod in a coastal marine protected area.*Canadian Journal of Fisheries and Aquatic Sciences* , 75, 1773-2108. doi: https://doi.org/10.1139/cjfas-2017-0254

Somero, G. N. (2002). Thermal physiology and vertical zonation of intertidal animals: optima, limits, and costs of living.*Integrative and Comparative Biology* , 42, 780–789. https://doi.org/10.1111/j.1365-2486.2005.00918.x

Southward, A. J., & Crisp, D. J. (1954). The distribution of certain intertidal animals around the Irish coast. *Proceedings of the Royal Irish Academy. Section B: Biological, Geological, and Chemical Science* , 57, 1-29. Retrieved August 11, 2020, from https://www.jstor.org/stable/20518835

Southward, A. J., Hawkins, S. J., & Burrows, M. T. (1995). Seventy years' observations of changes in distribution and abundance of zooplankton and intertidal organisms in the western English Channel in relation to rising sea temperature. *Journal of Thermal Biology* , 20, 127-155. doi: https://doi.org/10.1016/0306-4565(94)00043-I

Strathmann, M. F. (1987). *Reproduction and Development of Marine Invertebrates of the Northern Pacific Coast* . Seattle: University of Washington Press.

Thibert-Plante, X., & Hendry, A. P. (2010). When can ecological speciation be detected with neutral loci. *Molecular Ecology* , 19, 2301-2314. doi: 10.1111/j.1365-294X.2010.04641.x

Thorson, G. (1950). Reproductive and larval ecology of marine bottom invertebrates. *Biological Reviews* , 25, 1-45. doi: https://doi.org/10.1111/j.1469-185X.1950.tb00585.x

Tigano, A., & Friesen, V. L. (2016). Genomics of local adaptation with gene flow. *Molecular Ecology* , 25, 2144-2164. doi: https://doi.org/10.1111/mec.13606

Teske, P. R., Sandoval-Castillo, J., Golla, T. R., Emami-Khoyi, A., Tine, M., von der Heyden, S., & Beheregaray, L. B. (2019). Thermal selection as a driver of marine ecological speciation. *Proceedings of the Royal Society B: Biological Sciences* , 286, 20182023. doi: http://dx.doi.org/10.1098/rspb.2018.2023

Tomanek, L., & Helmuth, B. (2002). Physiological ecology of rocky intertidal organisms: a synergy of concepts. *Integraptive and Comparative Biology* , 42, 771-775. doi: https://doi.org/10.1093/icb/42.4.771

Treml, E. A., Roberts, J., Halpin, P. N., Possingham, H. P., & Riginos, C. (2015a). The emergent geography of biophysical dispersal barriers across the Indo-West Pacific. *Diversity and Distributions* , 21, 465-476. doi: 10.1111/ddi.12307

Underwood, A. J. (1972). Observations on the reproductive cycles of*Monodonta lineata* , *Gibbula umbilicalis* and *G. cineraria* . *Marine Biology* , 17, 333-340. doi: https://doi.org/10.1007/BF00366744

The UniProt Consortium. (2021). UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Research* , 49, D480-D489. doi: https://doi.org/10.1093/nar/gkaa1100

Waples, R. S. (1998). Separating the wheat from the chaff: patterns of genetic differentiation in high gene flow species. *Journal of Heredity* , 89, 438-450. doi: https://doi.org/10.1093/jhered/89.5.438

Wares, J. P. & Cunningham C. W. (2001). Phylogeography and historical ecology of the North Atlantic intertidal. *Evolution* , 55, 2455-2469. doi: https://doi.org/10.1111/j.0014-3820.2001.tb00760.x

Weersing, K., & Toonen, R. J. (2009). Population genetics, larval dispersal, and connectivity in marine systems. *Marine Ecology Progress Series* , 393, 1-12. doi: 10.3354/meps08287

White, C., Selkoe, K. A., Watson, J., Siegel, D. A., Zacherl, D. C., & Toonen, R. J. (2010). Ocean currents help explain population genetic structure. *Proceedings of the Royal Society B: Biological Sciences* , 277, 1685-1694. doi: 10.1098/rspb.2009.2214

Whitlock, M. C., & Lotterhos, K. E. (2015). Reliable detection of loci responsible for local adaptation: Inference of a neutral model through trimming the distribution of FST. *The American Naturalist* , 186, S24-S36. doi: 10.1086/682949

Williams, E. E. (1964). The growth and distribution of *Gibbula umbilicalis* (da Costa) on the rocky shore in Wales. *Journal of Animal Ecology* , 33, 433-442. doi: https://doi.org/10.2307/2563

Wort, E. J. G., Chapman, M. A., Hawkins, S. J., Henshall, L., Pita, A., Rius, M., ... Fenberg, P. B. (2019). Contrasting genetic structure of sympatric congeneric gastropods: do differences in habitat preference, abundance and distribution matter? *Journal of Biogeography* , 46, 369-380. doi: https://doi.org/10.1111/jbi.13502

Xuereb, A., Kimber, C. M., Curtis, J. M., Bernatchez, L., & Fortin, M-J. (2018). Putatively adaptive genetic variation in the giant California sea cucumber (*Parastichopus californicus*) as revealed by environmental association analysis of restriction-site associated DNA sequencing data. *Molecular Ecology* , 27, 5035-5048. doi: 10.1111/mec.14942

Xuereb, A., D'Aloia, C. C., Daigle, R. M., Andrello, M., Dalongeville, A., Manel, S., ... Fortin, M-J. (2019). Marine conservation and marine protected areas. In: Oleksiak, M., & Rajora, O. P. (Eds.),*Population Genomics: Marine Organisms* . (pp. 23). Switzerland: Springer Nature.

Yamada, S. B. (1989). Are direct developers more locally adapted than planktonic developers? *Marine Biology* , 103, 403-411. doi: https://doi.org/10.1007/BF00397275

Table 1. Number of polymorphic SNPs retained after each stage of the data filtering process in 2 RADseq datasets of the intertidal marine gastropods *Steromphala umbilicalis* and *Nucella lapillus* .

Filter	**# SNPs remaining after filtering**	**# SNPs remaining after filtering**
	S. umbilicalis	N. lapillus
Stacks catalogue	2,494,940	2,325,766
SNPs in >75% populations	2,494,940	2,325,766
SNPs with mean depth >10	1,438,613	154,624
SNPs with <20% missing data	121,613	29,287
SNPs with average MAF >0.05	15,682	5,529
SNPs with LD $r^2 < 0.7$	11,601	3,856
SNPs in HWE in at least 60% populations	11,600	3,844

Table 2. Descriptions of significant environmental (env) and spatial (dbMEM, AEM) vectors used in RDA analysis of *Steromphalaumbilicalis* and *Nucella lapillus* in the UK and Ireland.

Variable	**Description**
env_PC1	Winter air temperature (mean winter air temperature, minimum air temperature of coldest month)

Variable	Description
env_PC2	Summer air and sea surface temperature (mean and minimum summer sea surface temperature, maximum sum
env_PC3	Wave exposure and annual temperature range
dbMEM_2	Large-scale spatial pattern contrasting the North Channel/northern Irish Sea and north Wales coastline
dbMEM_3	Large-scale spatial pattern contrasting British (particularly sites in the south of Wales) from Irish sites
AEM5	Larval connectivity vector contrasting Selker Bay (in Liverpool Bay) with Barry (in the Bristol Channel)

Table 3. Number of outlier loci detected with six outlier detection methods in RADseq datasets of two intertidal marine gastropods, *Steromphala umbilicalis* and *Nucella lapillus* . Bayescan results are divided into outliers suggested to be under divergent and balancing selection. Brackets contain the number of outlier loci with Bayes Factors >100, indicative of "decisive" evidence of selection. Below, we provide the number of convergent outlier loci identified across an increasing number of methods.

Method of Outlier Detection (version)	Source	Cut-off Threshold	Number of Ou
			S. umbilicalis
PCAdapt (v4.3.3)	Luu et al. (2017)	Q<0.05	163
OutFLANK (v0.2)	Whitlock & Lotterhos (2015)	Q<0.05	115
BayeScan – divergent selection (v2.1)	Foll & Gaggiotti (2008)	Q<0.05	227 (143)
BayeScan – balancing selection (v2.1)	Foll & Gaggiotti (2008)	Q<0.05	0
Arlequin (v3.5.2.2)	Excoffier & Lischer (2010)	P<0.05	1,284
BayeScEnv (v1.1)	de Villemereuil & Gaggiotti (2015)	Q<0.05	98
LFMM (R package LEA v2.8.0)	Frichot & François (2015)	Q<0.05	17
Number of convergent methods			
1			1,425
2			113
3			77
4			52
5			14
6			0

Table 4. Results of full and partial redundancy analyses on neutral and outlier SNP datasets of two intertidal marine gastropods*Nucella lapillus* and *Steromphala umbilicalis* in the UK and Ireland. Models describe the proportion of genetic data (neutral or outlier SNPs) that is explained by environmental (env), spatial (dbMEM), and larval dispersal (AEM) predictor variables. The Anova(model) column gives the significance of the overall model estimated by permutation (only results for significant axes are shown). Partial RDAs are indicated by "pRDA" and the corresponding models are provided where conditioned predictors are given in brackets (e.g., *˜ dbMEM_2 + (env_PC1 + env_PC2)* indicates a partial RDA conducted on the geographic predictor dbMEM_2 after conditioning the model to remove the effects of environmental predictors env_PC1 and env_PC2).

Model	R^2adj	Anova(model) p-value
Nucella lapillus – **neutral loci**		
Best model ˜ env_PC1 + env_PC2	0.1862	0.0005
Nucella lapillus – **outlier loci**		
Best model ˜ dbMEM_2 + env_PC1 + env_PC2 + env_PC3	0.5756	0.0001

Model	**R^2adj**	**Anova(model) p-value**
Best model (pRDA) ~ dbMEM_2 + (env_PC1 + env_PC2 + env_PC3)	0.0004	0.4369
Best model (pRDA) ~ env_PC1 + env_PC2 + env_PC3 + (dbMEM_2)	0.2975	0.0022
Steromphala umbilicalis – **neutral loci**		
Best model ~AEM5 + dbMEM_2 + dbMEM_3 + env_PC2 + env_PC3	0.1739	0.0013
Best model (pRDA) ~AEM5 + (dbMEM_2 + dbMEM_3 + env_PC2 + env_PC3)	0.0281	0.2874
Best model (pRDA) ~ dbMEM_2 + dbMEM_3 + (AEM5 + env_PC2 + env_PC3)	0.0422	0.2020
Best model (pRDA) ~ env_PC2 + env_PC3 + (AEM5 + dbMEM_2 + dbMEM_3)	0.0207	0.2969
Steromphala umbilicalis – **outlier loci**		
Best model ~AEM5 + dbMEM_2 + dbMEM_3	0.4559	0.0009
Best model (pRDA) ~AEM5 + (dbMEM_2 + dbMEM_3)	0.0459	0.1353
Best model (pRDA) ~ dbMEM_2 + dbMEM_3 + (AEM5)	0.3940	0.0008

Table 5. BLAST matches for *Steromphala umbilicalis*(N=143) and *Nucella lapillus* (N=24) outlier loci contained within characterized coding regions.

Species	**Outlier Locus ID**	**Query Length (bp)**	**Sequence Match**	**Query Cover (%)**	**E-value**	**Percent Identity (%)**	**Accession Number**
S. umbilicalis	39885	226	*Nucella lapillus* estrogen receptor mRNA, complete cds	34	3.00E-11	83.12	EF591073.1

Species	Outlier Locus ID	Query Length (bp)	Sequence Match	Query Cover (%)	E-value	Percent Identity (%)	Accession Number
	41190	290	PREDICTED: *Arvicanthis niloticus* islet cell autoantigen 1 like (Ica1l), transcript variant X1, mRNA PREDICTED: *Arvicanthis niloticus* islet cell autoantigen 1 like (Ica1l), transcript variant X2, mRNA	44	8.00E-21	80.62	XM_034499185.1; XM_034499186.1
	88303	237	PREDICTED: *Aplysia californica* cilia- and flagella-associated protein 100 (LOC101850599), transcript variant X1, mRNA PREDICTED: *Aplysia californica* cilia- and flagella-associated protein 100 (LOC101850599), transcript variant X2, mRNA	41	9.00E-19	83.67	XM_005108065.3; XM_013087902.2

Species	Outlier Locus ID	Query Length (bp)	Sequence Match	Query Cover (%)	E-value	Percent Identity (%)	Accession Number
	2419290	271	PREDICTED: *Arvicanthis niloticus* islet cell autoantigen 1 like (Ica1l), transcript variant X1, mRNA PREDICTED: *Arvicanthis niloticus* islet cell autoantigen 1 like (Ica1l), transcript variant X2, mRNA	39	1.00E-12	78.9	XM_034499185.1; XM_034499186.1
N. lapillus	3967	291	*Nucella lapillus* estrogen receptor gene, partial cds	35	9.00E-14	78.85	EF591072.1
	35726	290	*Nucella lapillus* estrogen receptor gene, partial cds	44	3.00E-26	82.01	EF591072.1
	42831	330	*Pomacea canaliculata* uncharacterized protein	65	2.40E-27	62.2	A0A2T7PAW8

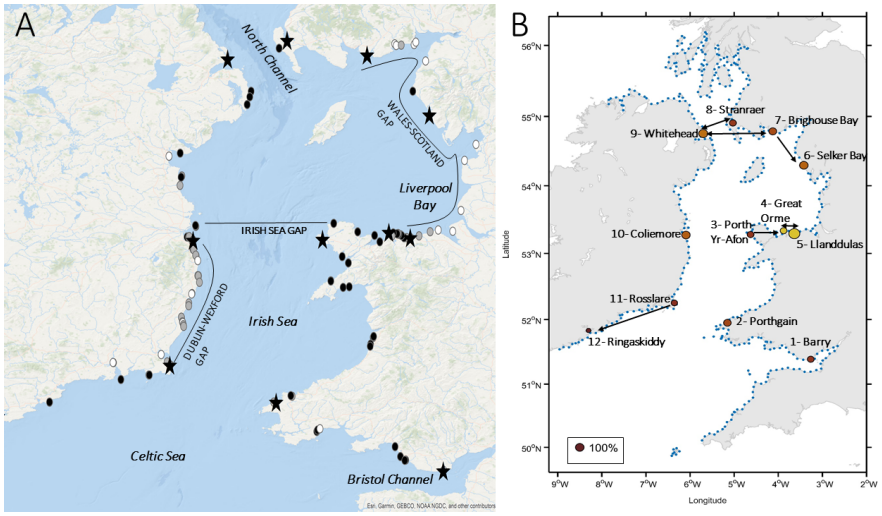

Figure 1. (A) Sites surveyed (circles) and sampled (stars) in the UK and Ireland for the intertidal marine invertebrates *Nucella lapillus*and *Steromphala umbilicalis* . Sites are coloured according to the presence or absence of one or both species (black = both species present; grey = only *N. lapillus* present; white = both species absent). Identified habitat "gaps" are indicated with black lines and labelled in capitals. Water bodies are labelled in italics. (B) Release locations of the gastropod *Steromphala umbilicalis* , in the larval dispersal model, showing sample sites (labelled, red-yellow points) and interim stepping-stones sites (blue points). The size and colour of sample site markers indicate the degree of convergence (i.e., net sink sites are large and yellow) or divergence (i.e., net source sites are small and dark red) for one season of dispersal (June-September). The "100%" legend in the bottom left corner indicates the size of a site marker that had no net gain or loss of particles at the end of a spawning season. The black arrows indicate the sample sites which are connected (no year-on-year spread shown, background connectivity only). Online version in colour.

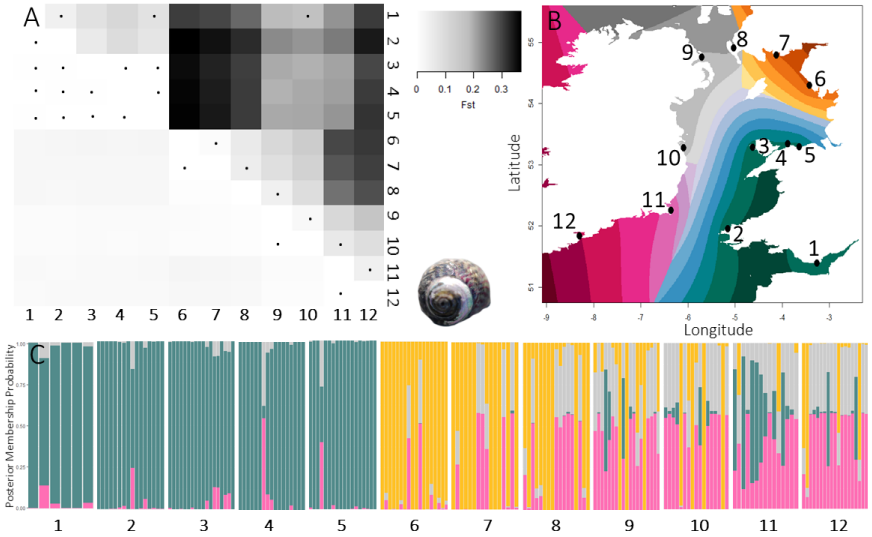

Figure 2. Population structure analysis of 11,457 neutral and 143 outlier SNPs for the intertidal gastropod *Steromphala umbilicalis* in the UK and Ireland. (A) Heatmap based on pairwise F_{ST} values across sampling sites (neutral SNPs = below diagonal; outlier SNPs = above diagonal). Non-significant pairwise comparisons

are indicated with a black dot (P>0.001). The distribution of population structure for K=4 adaptive genetic clusters (outlier SNPs) is shown using TESS3 (B) and DAPC (C). Online version in colour.

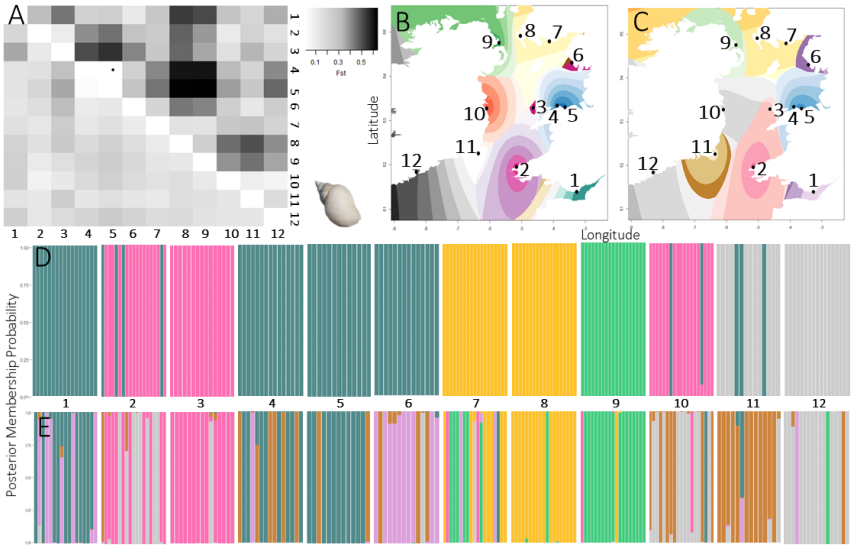

Figure 3. Population structure analysis of 3,820 neutral and 24 outlier SNPs for the intertidal gastropod *Nucella lapillus* in the UK and Ireland. (A) Heatmap based on pairwise F_{ST} values across sampling sites (neutral SNPs = below diagonal; outlier SNPs = above diagonal). Non-significant pairwise comparisons are indicated with a black dot (P>0.001). The distribution of population structure for K=5 (D) and K=11 (B) neutral genetic clusters, and K=7 (E) and K=8 (C) adaptive genetic clusters identified using TESS3 (B-C), and DAPC (D-E). Online version in colour.

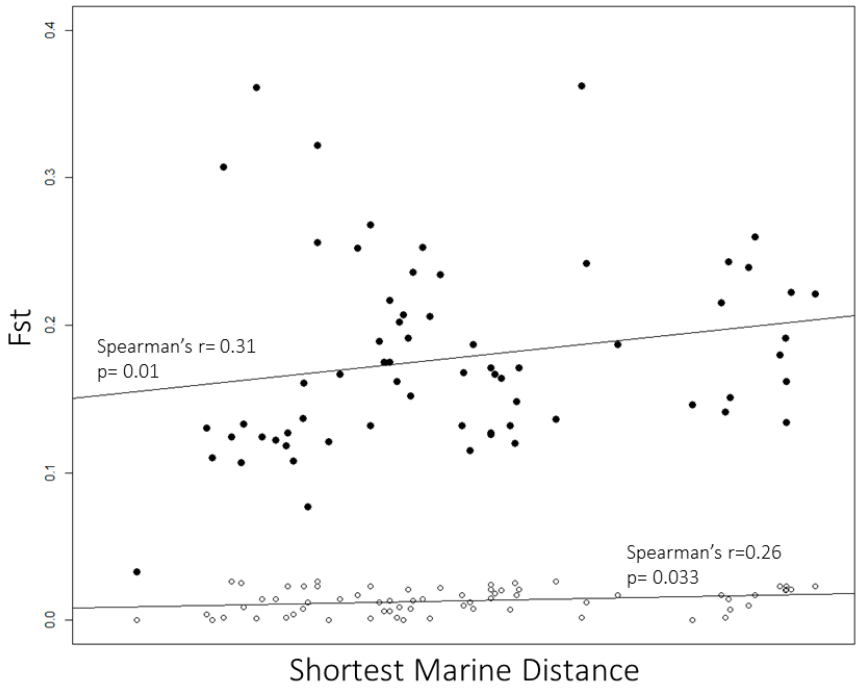

Figure 4. Relationship between geographic (shortest marine distance) and genetic (neutral F_{ST}) distances between pairwise population comparisons of the intertidal marine gastropods*Steromphala umbilicalis* (open circles) and *Nucella lapillus* (closed circles). Spearman's rank correlation coefficients were used to test the significance of the linear relationships.

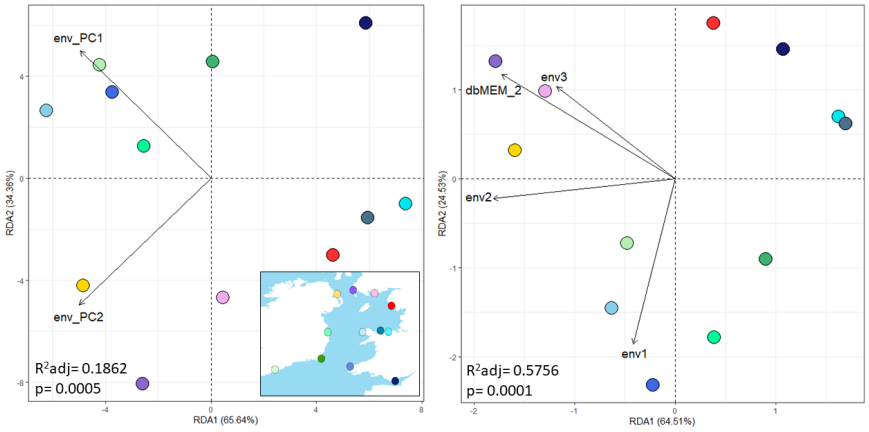

Figure 5. Biplots showing results of RDA analyses on the best model for neutral (left, N= 3,820) and outlier (right, N= 24) SNP datasets of the intertidal marine gastropod *Nucella lapillus* . Coloured points on the biplots correspond to sampling sites indicated on the inset map. Arrows represent predictor variables that significantly explain population genetic structure for each dataset. Online version in colour.

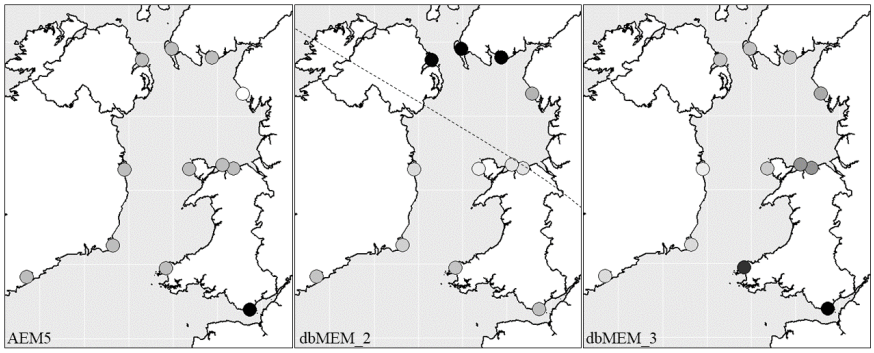

Figure 6. Plot of significant dbMEM and AEM vectors identified in seascape genomic RDA analyses of *Steromphala umbilicalis* and*Nucella lapillus* in the UK and Ireland. Symbols represent sampled sites for this study, and colours correspond to similarity in eigenvector values (black=high, grey=mid, white=low). The dbMEMs are associated with large-scale spatial structure in the data, given by their large eigenvalues (see also Figure S1). The dashed line in dbMEM_2 corresponds to the relative positioning of the Forbes line, demarcating the general northern limit of southern species within the UK and Ireland (Forbes 1858).

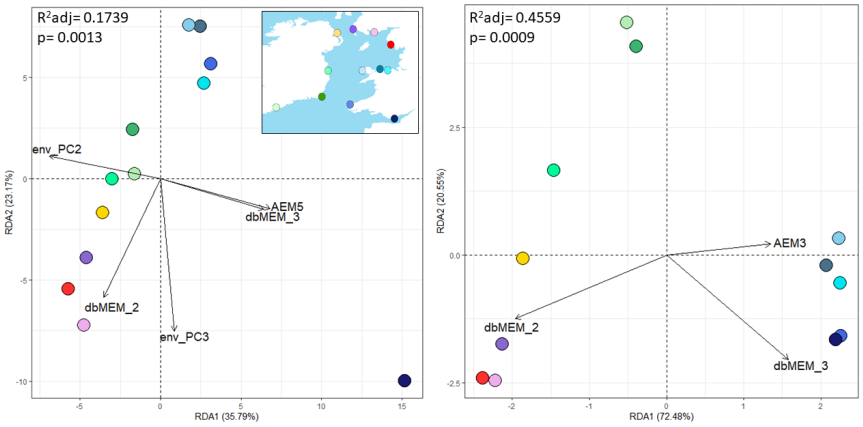

Figure 7. Biplots showing results of RDA analyses on the best model for neutral (left, N= 11,457) and outlier (right, N= 143) SNP datasets of the intertidal marine gastropod *Steromphala umbilicalis* . Coloured points on the biplots correspond to sampling sites indicated on the inset map. Arrows represent predictor variables that significantly explain population genetic structure for each dataset. Online version in colour.