



RESEARCH ARTICLE

Seascape genetics in a polychaete worm: Disentangling the roles of a biogeographic barrier and environmental factors

Cecili B. Mendes¹ | Thainá Cortez¹ | Cinthya S. G. Santos² |
Thadeu Sobral-Souza³ | Arian Dielactaquiz Santos⁴ | Dalton Kei Sasaki⁴ |
Danilo Augusto Silva⁴ | Marcelo Dottori⁴ | Sônia C. S. Andrade¹

¹Laboratório de Diversidade Genômica, Departamento de Genética e Biologia Evolutiva do Instituto de Biociências, Universidade de São Paulo (USP), São Paulo, Brazil

²Laboratório de Sistemática e Ecologia de Polychaeta, Departamento de Biologia Marinha, Universidade Federal Fluminense (UFF), Niterói, Brazil

³Macroecology Lab (MacrEco), Departamento de Botânica e Ecologia, Universidade Federal de Mato grosso (UFMT), Cuiabá, Brazil

⁴Coastal Hydrodynamics Laboratory (LHiCo), Instituto Oceanográfico, Universidade de São Paulo (USP), São Paulo, Brazil

Correspondence

Sônia C. S. Andrade, Laboratório de Diversidade Genômica, Departamento de Genética e Biologia Evolutiva do Instituto de Biociências, Universidade de São Paulo (USP), São Paulo, Brazil.
Email: soniacsandrade@ib.usp.br

Funding information

Coordenação de Aperfeiçoamento de Pessoal de Nível Superior; Fundação de Amparo à Pesquisa do Estado de São Paulo, Grant/Award Number: 2015/20139-9 and 2016/20005-5

Handling Editor: Greer Dolby

Abstract

Aim: Seascape genomic studies aim to understand how environmental variables shape species diversity through genotype–environment associations. Identifying these effects on lecithotrophic larval species that live in intertidal zones is particularly challenging because they are subject to environmental heterogeneity and anthropogenic events. Here, we evaluate how biotic and abiotic features in the Southwest Atlantic littoral zone can affect a high dispersal species' present and historical demography.

Location: Brazilian coast, covering more than 3000 km.

Taxon: *Perinereis ponteni*.

Methods: We investigated population genetic diversity, connectivity and past dynamics using 23,300 single-nucleotide polymorphisms (SNPs) generated using genotyping by sequencing. We tested whether environmental abiotic variables could explain the variance found in genotype frequencies using isolation-by-environment (IBE) and landscape association approaches. These data, combined with palaeodistribution simulations and oceanic circulation modelling, were used to infer species demographic history and connectivity patterns.

Results: Along with high levels of connectivity detected, we found a genetic boundary in the southeastern region of Brazil around Cabo Frio (Rio de Janeiro), and a cline trend for some loci. The palaeodistribution simulations reveal a spatial refuge in the southeast during the Last Glacial Maximum (21 kya), with the expansion of the northern region. We identified 1421 SNPs with frequencies associated with eight environmental variables, most of which were related to temperature—the main environmental factor determining IBE.

Main Conclusions: *Perinereis ponteni*, a polychaete with high gene flow capability responds to biogeographic barriers, highlighting the importance of biotic and abiotic factors in shaping population connectivity. Furthermore, the effect of temperature indicates that future climate change and ocean warming can hugely impact this species.

KEYWORDS

ecological niche modelling, genetic diversity, Nereididae, ocean currents, *Perinereis ponteni*, rocky shores

1 | INTRODUCTION

Marine populations were long thought to be panmictic, with high rates of gene flow (Caley et al., 1996; Palumbi, 1994), explained by the lack of explicit geographical barriers. Some barriers to gene flow in marine systems are seasonally determined, resulting in partially isolated populations (Palumbi, 1994). In most marine species, the larval phase is the dominant dispersal stage, and pelagic larval duration (PLD) is directly linked to dispersal ability (Bohonak, 1999; Levin, 2006; Pascual et al., 2017). Consequently, species with planktonic larval stages are more likely to present better connected populations than those without planktonic stages. Nevertheless, many studies have shown that complex interactions between species-specific traits and environmental characteristics result in different patterns of connectivity that cannot be explained solely by larval dispersion through ocean currents (Bernardi et al., 1993; Bernatchez et al., 2018). Environmental conditions like temperature and salinity are suggested to have adaptive importance in shaping gene flow in marine species (Bernatchez et al., 2018; Saenz-Agudelo et al., 2015; Sandoval-Castillo et al., 2017).

Seascape genetic studies of single-nucleotide polymorphisms (SNPs) are highly sensitive for detecting patterns of population isolation and local adaptation because of the high number of markers

analysed. A few hundred SNPs can detect discrete changes in gene flow (Bernatchez et al., 2018; Saenz-Agudelo et al., 2015; Selkoe et al., 2016), allowing the identification of genomic regions that are possibly involved in the adaptive process, which provides important insights into a changing environment. Among invertebrates, most studies have focused on species of commercial importance, primarily molluscs and crustaceans. Therefore, these patterns remain unknown for many invertebrate groups, such as polychaetes, which are crucial in marine communities. Polychaeta includes diverse invertebrates under Annelida, comprising burial, planktonic, tubedwelling and reef-building species (Nunes et al., 2021; Pamungkas et al., 2019). They are among the most abundant and species-rich marine metazoans in benthic environments (Díaz-Castañeda & Reish, 2009; Fauchald & Jumars, 1979; Rouse & Pleijel, 2001). *Perinereis ponteni* Kinberg, 1865, is a Nereididae polychaete species with a broad spatial distribution in the Atlantic Ocean, from Mexico to Brazil (Figure 1). Although these animals are more easily found in mussels of the *Brachidontes* genus, they also inhabit algae, oysters and barnacle beds, acting as essential food generalists in these communities. Despite their ecological importance, there have been no detailed studies of their reproduction. The species is known to exhibit epitoky with a planktonic stage and might have a lecithotrophic larval phase like most Nereididae species (Bakken et al., 2018;

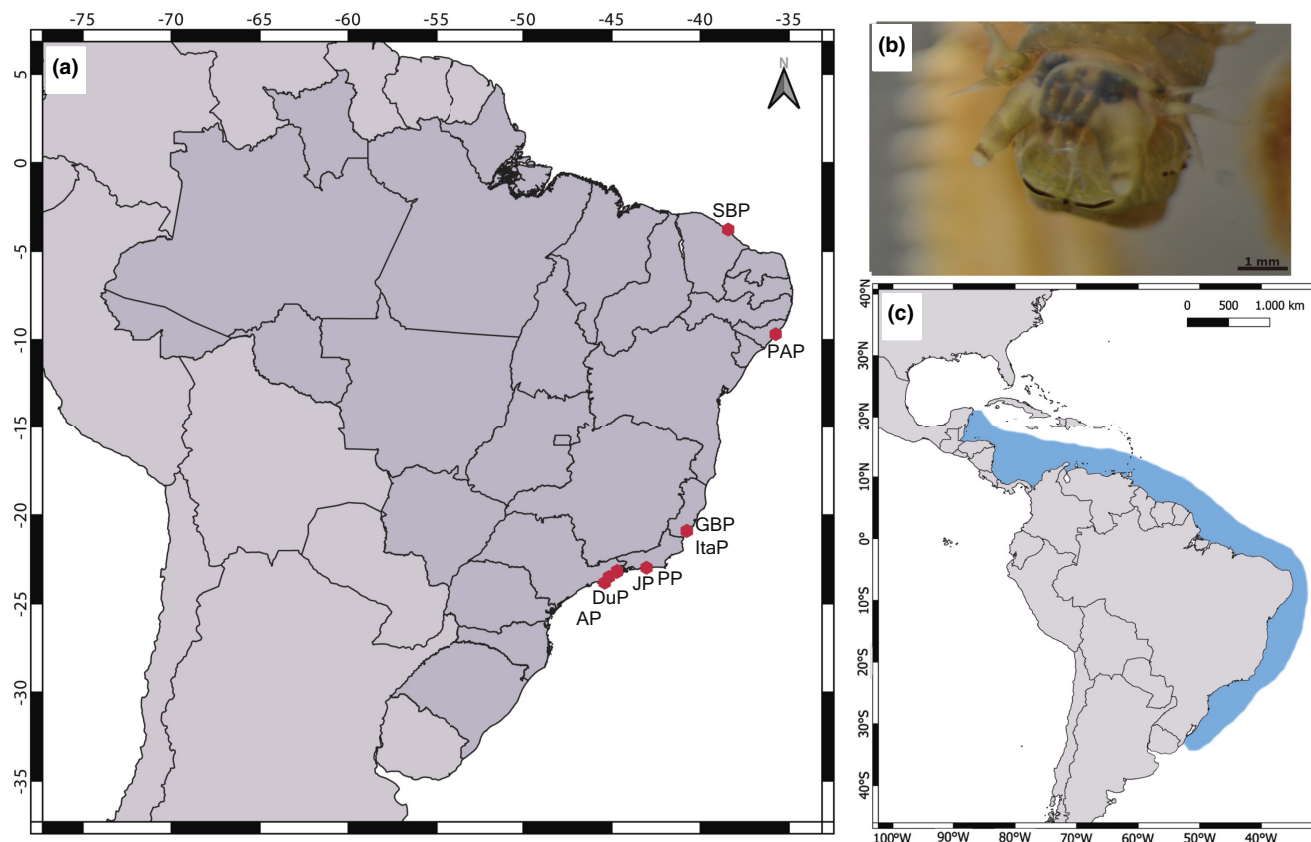


FIGURE 1 (a) *Perinereis ponteni* collection sites from North to South—Sabiaguaba (SBP); Prado (PAP); Gamboa (GBP); Itaoca (ItaP); Praia (PP); Jabaquara (JP); Praia dura (DuP); Araçá (AP); (b) *P. ponteni* head detail; (c) known distribution, from Mexico to South Brazil. The projection of the main map is the Equirectangular.



Rouse, 2000). Previous studies using mitochondrial markers to investigate the population structure of *P. ponteni* on the Brazilian coast found a high level of mixture, with high rates of gene flow among all populations (Paiva et al., 2019).

In this study, high levels of gene flow between the Brazilian populations of *P. ponteni* were expected. Any disruption in this pattern is likely a result of environmental factors possibly acting on the larval dispersion and post-settlement survival rates as well as the past demographic scenarios. This study is among the first to evaluate the environmental impact on population connectivity of an important invertebrate species in the Southwest Atlantic Ocean. We coupled genomic scale sequencing with ecological association tests, oceanic circulation and niche-based models to understand the present and historic gene flow patterns and how these can be affected by a changing environment.

2 | MATERIALS AND METHODS

2.1 | Sampling

The specimens were collected from oyster and barnacle clusters in eight localities along the Brazilian coast (Figure 1; Supplementary Material) from 2017 to 2019. The clusters were collected from exposed rocks and pillars with a spatula and transported to the laboratory in plastic bags. In the laboratory, clusters were submerged in seawater for a few hours before being taken apart. Subsequently, all worms were carefully removed with forceps. The specimens were relaxed using pulverized menthol dissolved in seawater and identified under a stereomicroscope. Subsequently, the worms were fixed in 99% ethanol and stored at -20°C until DNA extraction.

2.2 | DNA extraction, library preparation and sequencing

Genomic DNA was extracted from 61 specimens of *P. ponteni* using the protocol by Doyle and Doyle (1987). DNA quality and concentration were assessed by gel electrophoresis, spectrometry (NanoDrop Lite spectrophotometer; Thermo Fisher Scientific) and fluorometry (Qubit 3.0 Fluorometer; Invitrogen). Approximately 100 ng of the extracted DNA per sample was sent to EcoMol Consultoria e Projetos Ltd., where GBS libraries were prepared according to the protocol described by Elshire et al. (2011) and modified by Nunes et al. (2017). The libraries were sequenced in 100 bp single-end fragments using the Illumina HiSeq 2500 platform at the Centro de Genômica Funcional ESALQ-USP.

2.3 | SNPs prospecting and data analysis

Sequenced fragments were first filtered using Seqclean v1.10.09 (Zhbannikov et al., 2017), removing sequences with an average

Phred quality score ≤ 20 , contaminants and adapters (UniVec database; <https://www.ncbi.nlm.nih.gov/tools/vecscreen/univec/>). After filtering, Ipyrad v0.7.30 (Eaton & Overcast, 2020) was used for demultiplexing, quality and size filtering, fragment clustering and SNP prospecting. Reads with more than five Ns or shorter than 35 bp were discarded. The minimum read depth was set to six for calling consensus sequences within the samples, and the maximum depth was set to 10,000. The clustering threshold was set to 90%, and the maximum number of SNPs per locus was set to 30. A locus had to be present in 50% or more of the samples in order to be retained in the final dataset. All other parameters were maintained at their default values. SNPs were first visualized as an occupancy matrix using the divergent option in the Matrix Condenser tool (de Medeiros & Farrell, 2018); samples with $\geq 40\%$ missing data were removed. The resulting variant call format (VCF) files were used as input to VCF tools 0.1.16, with the option max-missing 0.5 (Danecek, et al., 2011), followed by the PLINK software (Purcell et al., 2007), which filtered SNPs for missing data (geno < 0.45), rare alleles (minor allele frequency > 0.01) and linkage disequilibrium (indep-pairwise 50 5 0.5). The resulting files were converted to specific program formats using PGDSpider 2.1.1.5 (Lischer & Excoffier, 2012).

2.4 | Population structure

To characterize each population, molecular diversity indices (nucleotide diversity $[\theta_{\pi}]$ and nucleotide differences $[\theta_{\xi}]$) were calculated using Arlequin 3.5.2.2 (Excoffier & Lischer, 2010) using 1000 simulations. Arlequin was also used to calculate pairwise F_{ST} , global and population-specific F_{IS} (Weir & Cockerham, 1984), and analysis of molecular variance (AMOVA). The multiloci estimates of observed (H_O) and expected (H_E) heterozygosity and global F_{ST} (Weir & Cockerham, 1984) were calculated using the 'adegenet' package (Jombart, 2008) in R (R Core Team, 2013).

Structure software (Pritchard et al., 2000) was used to analyse the complete SNP dataset, putatively outliers and neutral SNPs (see below), using values of K from 1 to 16 with default settings, except for lambda and alpha values. The alpha value was set to one in all runs, and the lambda value was first estimated in one run with all other values fixed. The estimated lambda value was fixed for 10 other runs. The results from the 10 iterations were compacted and analysed using the R package Structure Harvester (Earl, 2012), where the likelihood of each K was compared, and the most suitable K was selected. The selected K was used to generate a bar plot showing the most likely ancestry of each individual using the R package 'PopHelper' (Francis, 2017).

The dataset was used in the 'vegan' package (Oksanen et al., 2020) as input for the Mantel test (Mantel, 1967) to identify possible isolation-by-distance (IBD), and for the isolation-by-environment (IBE) analysis (Wang & Bradburd, 2014). Each collection site was considered a different population for this test. The genetic distance was inferred using pairwise linearized F_{ST} (Weir & Cockerham, 1984). The smallest distance between the two sites

was calculated as the line along the coast connecting the two points using Google Maps (Google Corporation). Mantel and partial Mantel tests were performed with 1000 simulations. Partial Mantel tests were performed separately for each environmental variable using geographic distance as a control for 18 variables (Table S1). The present and past conditions from the mid-Holocene were considered for each variable because they were the same as those used for the palaeodistribution simulations (see below and Table S1, first 18 rows). Climatic distances were calculated as Euclidean distances using the `dist` function in R.

2.5 | Cline analyses

The genetic transition between populations through cline-fitting models was performed in 'hzar' (Derryberry et al., 2014). This R package allows flexibility in fitting cline models of varying complexity owing to its modular design. This package uses the allele frequencies of each SNP and attempts to correlate its frequency with geographic distribution, creating a cline-fitting model for each allele. Only 100 SNPs with a higher contribution to principal component analysis (PCA) structuring were used (Figure S1), as they had the most contrasting frequencies among the populations. Sixteen models, including a null model, were tested by combining the different tails and scaling values for each SNP (Table S2). Three independent runs for each model were conducted, each consisting of 20,000 iterations, following an initial burn-in phase of 500 iterations. The model with the lowest Akaike information criterion (AIC) for each locus was retained. Only the results where the null model was not accepted are discussed.

2.6 | Demographic analyses and estimates of historical migration

The migration matrix and demographic scenarios were estimated using *Fastsimcoal* 2.7 (Excoffier & Foll, 2011). The minor allele site frequency spectrum parameter computed with *Arlequin* with 1000 bootstraps was used as the input. In the population expansion scenarios, the two clusters identified in the *Structure* software were used as populations, and the two models were tested. The first model estimated population expansion from the northeast (NE), with posterior migration towards the southeast (SE). The second considered population expansion from the SE, with posterior northward migration. Fifty independent replicates were performed for each model. Each of these replicates included 40 estimation loops with 300,000 coalescence simulations. Given the observed data, the probability of each model was determined based on the maximum likelihood value and AIC. The historical migration rates were estimated using each locality as a population, with 10 independent replicates, each including 40 estimation loops with 60,000 coalescence simulations, and assuming current migration between all pairs of populations. In both cases, the

mutation rate was also estimated using *Fastsimcoal*, with the first cluster's results ranging from $1e^{-9}$ to $1e^{-6}$.

The population connection through larval dispersion was simulated using *OpenDrift* (1.7.1, Dagestad et al., 2017), an open-source particle tracking framework that uses the Eulerian velocity and thermohaline fields from an Eulerian model and a fourth-order Runge-Kutta (Edwards et al., 2006) scheme to transport particles within the domain. The *OpenDrift* module *Larvalfish* has been adapted to *P. ponteni* larvae, where the functions of fish growth (Folkvord, 2005) and vertical migration of larvae were applied to fit the module to the maximum length of the organism ($\sim 310\mu\text{m}$) and a maximum depth of vertical migration ($\sim 2.5\text{m}$). The adapted module considers the drift composed of ocean currents, wind drift at the surface, vertical mixing, vertical migration, the hatching from an egg to larvae and its continuous growth and the terminal velocity, according to Sundby (1983).

The *OpenDrift* simulations consisted of 1 year (from 1 January 2019 to 2 January 2020), with daily releases of 1000 particles (i.e. eggs) inside a 10-m radius at Espírito Santo (41°W , 22°S), Ubatuba (45.1014°W , 23.5218°S) and São Sebastião (45.3687°W , 23.7406°S). A total of 1000 particles per site were used because of the trade-off between the statistical significance and computational cost of processing and analysis. Each simulation had a time-step of 10,800 seconds (3h) and a vertical mixing resolution of 60s. After 50 days of drifting, the larvae were considered dead if they were not stranded on the shore. The current fields were provided by the forecast system *PreAMar* (Costa et al., 2020), which simulates conservative variables (temperature, pressure and salinity), currents and sea level (Blumberg et al., 2010). The choice of the southwest coast of Brazil to simulate larval dispersion was due to two main factors: (i) a gene flow barrier occurs in this region and (ii) the distance from and between the other two data locations is considerable; therefore, the currents would not allow a displacement long enough to establish a direct connection between those populations.

2.7 | Outlier SNPs

To identify SNPs under selection, the SNP matrix was used as input in *BayeScan* 2.0 (Foll & Gaggiotti, 2008), using 20 pilot runs of 50,000 iterations followed by 100,000 simulations with prior odds of 100 and 1% false discovery rate. Each location was considered a different population.

The genotype-environment association method redundancy analysis (RDA) was applied to detect possible SNPs under selection (Forester et al., 2018). The environmental variables analysed were the same as those used in the IBE inferences (Table S1). First, a correlation analysis was performed using all 18 variables, and 11 highly correlated variables ($|r| > 0.7$) were removed. The remaining seven variables (annual mean sea surface salinity, annual mean sea surface temperature, bathymetry, east-west aspect, north-south aspect, profile curvature, and concavity) were set as RDA predictors, and only the SNPs indicated by the significant RDs were considered.

Possible associations between environmental variables and SNP frequency were also assessed with the latent factor mixed model tests (LFMM; Frichot et al., 2013; Rellstab et al., 2015) implemented in the R package 'LEA' (Frichot & François, 2015). To select the environmental variables, we first ran a PCA using 24 environmental variables (Table S3) from Bio-Oracle (<https://bio-oracle.org/>). We selected four variables showing the most substantial contribution to the first two PCA axes (chlorophyll A, carbon phytoplankton biomass, pH and nitrate concentration). We then performed a second PCA with 19 variables (Table S3) from WorldClim (www.worldclim.org). We selected the four variables showing the most substantial contribution to the first two PCA axes (temperature seasonality, minimum temperature of the coldest month, mean temperature of wettest quarter and mean temperature of the warmest quarter). These eight environmental variables (four from Bio-Oracle and four from WorldClim) were used to run the LFMM. To avoid Type I and II errors, which can arise from incorrect assumptions about the underlying demographic structure (Storfer et al., 2018), we ran LFMM using both the number of clusters defined by Structure and the number of sampling locations, each doubled, and retained only the SNPs resulting from all runs. The environmental variables used in LFMM and RDA were different owing to the type of analysis performed by each software. While RDA pools all variables during the analysis, LFMM deals with each variable separately. Therefore, certain environmental variables must be prioritized. To do so, we used two PCAs to choose those with higher effects on the species distribution.

Outlier SNPs identified by at least one software were blasted against NCBI's nt and nr databases. Annotated proteins were searched in the UniProt database (<https://www.uniprot.org>) to determine their functions.

2.8 | Niche-based inferences

The palaeodistribution of *P. ponteni* was based on niche-based models that use known occurrences and climate variables to predict species distribution (Alvarado-Serrano & Knowles, 2014; Ferrier & Guisan, 2006; Stigall, 2012). To compile the known species occurrences, we used our sampling information and searched the Web of Science® database (Institute of Scientific Information; Thomson Scientific) using the following keywords and Boolean command combinations: Perinereis, AND ponteni, AND distribution* occurrence* record. Consequently, 23 occurrence sites were compiled, encompassing the entire species distribution (Table S4). Initially, we used all ocean and geophysical variables available in the Marspec database (<https://www.marspec.org>) (Sbrocco, 2014; Sbrocco & Barber, 2013) for current and past climate scenarios separately with 5 arc-min cell size resolution of approximately 10 km × 10 km. These variables were derived from coupled ocean-atmosphere general circulation models available through the second phase of the Palaeoclimate Modelling Intercomparison Project (PMIP2). As these variables were correlated with others,

we conducted a factorial analysis to select noncorrelated variables. We used bathymetry, plan curvature, profile curvature, salinity of the saltiest month, annual salinity range and annual sea temperature range (°C) to build the niche-based models. Different algorithms are currently able to predict the species distribution (Araújo et al., 2019), and the combined use with ensemble forecasting is recommended. We used four algorithms corresponding to presence-only methods: (i) envelope score (bioclim), (ii) Gower distance and presence-background methods, (iii) MaxEnt and (iv) support vector machine (SVM). We used 75% presence to build the models and 25% for testing. We used the true skill statistic (TSS) value for model evaluation (Allouche et al., 2006) and the maximum sensitivity and specificity as thresholds (Table S5). All models were built for current climate conditions and projected to the Last Glacial Maximum (LGM, 21 kya) and Holocene (6 kya) climate conditions. All models were built using 'dismo', 'sf' and 'kernlab' packages in R environment.

3 | RESULTS

3.1 | Prospection of SNPs and genomic diversity

The sequenced libraries resulted in 253,624,109 reads (BioProject PRJNA750868), with reads per sample ranging from 607,246 to 13,351,934. After several filtering and clustering steps, the average number of consensus reads was 2,132,741 (Table S6). We analysed 23,300 SNPs from 61 specimens, with the SNPs per sample ranging from 13,777 to 23,296 and a median of 22,000 SNPs per individual. The SNPs revealed that the observed heterozygosity across all loci was significantly lower than expected for all localities (Table S7). The fixation coefficient F_{IS} values across all loci were significant in seven of the eight localities, ranging from 0.118 to 0.172 (Table S7).

3.2 | Population structure

The Structure results from the matrices with 23,300 SNPs apportioned to the specimens in two clusters ($\Delta K = 44.075$, likelihood = -588,114; Figure S2). The first encompassed all samples from the NE, Itaoca (SE), and Gamboa (SE). The second sample contained the remaining SE samples (Figure 2a). The PCA analyses showed a different pattern with four groups delimited: the first with NE populations, the second with populations from Espírito Santo, the third with specimens from Prainha and the fourth with the remaining SE populations (Figure 2b). The results of the pairwise F_{ST} showed significant values in most comparisons, ranging from 0.017 to 0.064 (Table S8). For AMOVA analyses, we defined each sampling site as a population and the two clusters defined by Structure as groups. The results revealed that most of the variation was within individuals (84.05%), followed by variation among individuals in each population (13.14%) (Table S9).

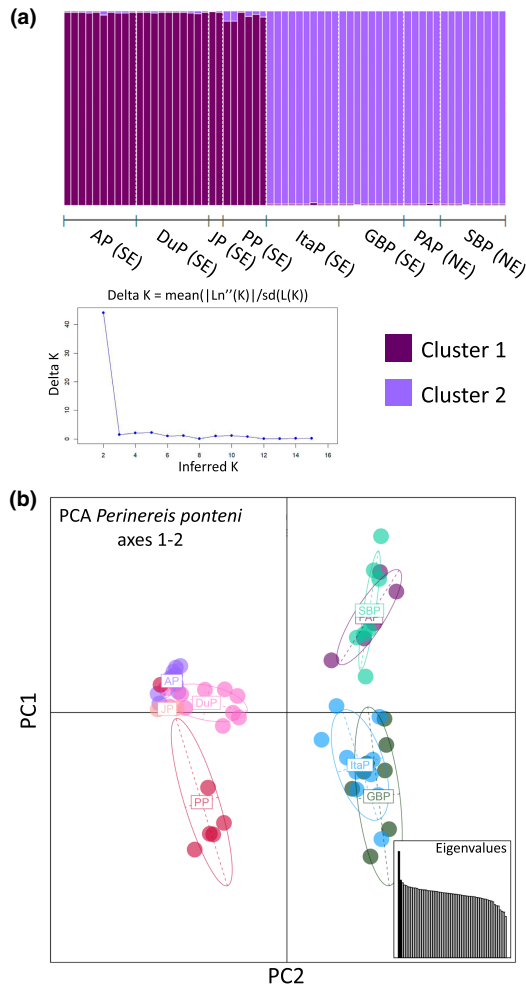


FIGURE 2 Clustering results for 23,300 SNPs from 61 *Perinereis ponteni* specimens from northeast (NE) and southeast (SE) regions. (a) Bar chart showing Structure results for $K = 2$ and Structure Harvester results indicating $K = 2$ as the most probable K value. Each bar represents an individual and each colour represents a different cluster. Horizontal lines delimit the sampling sites; (b) PCA results, PC1 (eigenvalue = 1276.3325; 3% of the variance) and PC2 (eigenvalue = 933.2545; 2% of the variance). Each dot corresponds to one specimen. Different colours represent different sampling sites. Araçá (AP), Praia dura (DuP), Jabaquara (JP), Prainha (PP), Itaoca (ItaP), Gamboa (GBP), Prado (PAP), Sabiaguaba (SBP). SNP, single-nucleotide polymorphism.

3.3 | Cline analyses

Of the 100 SNPs that contributed the most to genetic structuring, five were outlier SNPs, and 95 were putatively neutral SNPs. The null model had the lowest AIC values for all outlier SNPs and the 81 putatively neutral SNPs. Therefore, these SNPs cannot produce a cline-fitting graph. The Hzar test indicated a cline in the geographic area between the two clusters delimited by the Structure software. Fourteen putatively neutral SNPs presented the lowest AIC for Model 1 (no tail and fixed scaling) (Figure S3). The centre slope of these clines is approximately 2900 km from Sabiaguaba (NE), which

is near Prainha (SE). All clines showed soft slopes, ranging between 0.00024978 and 0.0135, with SNP 763 showing the sharpest cline (slope = 0.0135, width = 29.615). The mean ΔP (mean change in allele frequency) of the cline tail values was relatively high ($\Delta P = 0.65856$), indicating an evident differentiation.

3.4 | Demographic analyses and estimates of migration

The optimal-fitting model tested by Fastsimcoal, according to the highest likelihood (-6415.83) and the lowest AIC (29,258), was the one in which the population divergence initiated in the northeastern region towards the south approximately 555,000 generations ago. The best model assumed an initial unidirectional gene flow, followed by bidirectional migration (Figure 3). Migration rates showed a nearly symmetrical gene flow between NE and SE, with SE populations receiving slightly fewer migrants (Figure 4). The Mantel test showed no statistically significant correlation between genetic and geographic distances ($R = -0.0590.393$, $p = 0.500$), indicating no IBD. Partial Mantel tests indicated that sea surface temperature was a significant factor for IBE, both for the present ($R = 0.407$; $p = 0.022$) and past ($R = 0.414$; $p = 0.027$) populations (Table S10).

The larval dispersion simulation showed that most larvae stayed in the region where they were released, with only a few individuals transported to distances greater than 50 km from the seeding points (Figure 5). Individuals released in the vicinity of Itaoca (ES) seem to travel mostly southward. They are transported only to immediately after the Cabo Frio region (around Saquarema, RJ). In contrast, the larvae that pass this area cannot reach any part of the coast before the maximum time to metamorphosis (Figure 5a). The larvae released around Praia Dura (SP) and Araçá (SP), however, seemed to mostly travel northward, reaching the coast of Rio de Janeiro in the vicinity of Cabo Frio (Figure 5b,c). Nonetheless, the larvae released in this area (Araçá and Praia Dura), which travel southward after the São Sebastião channel, do not reach the coast in any area. This suggests that the Cabo Frio barrier is mainly determined by the hydrodynamic pattern (currents) rather than the thermodynamic pattern (temperature) associated with the local upwelling system.

3.5 | Outlier SNPs

BayeScan identified 12 outlier SNPs, all were indicative of positive selection (Figure S4). In contrast, LFMM indicated the presence of 404 loci (1421 SNPs) with allele frequencies correlated to the eight environmental variables tested (Table S11; Figures S5 and S6). RDA indicated 85 SNPs (Figure S7) related to seven variables (longitude, latitude, annual mean sea surface salinity, annual mean sea surface temperature, bathymetry, East-West aspect, North-South aspect, profile curvature and concavity). Six SNPs were identified using both RDA and LFMM,

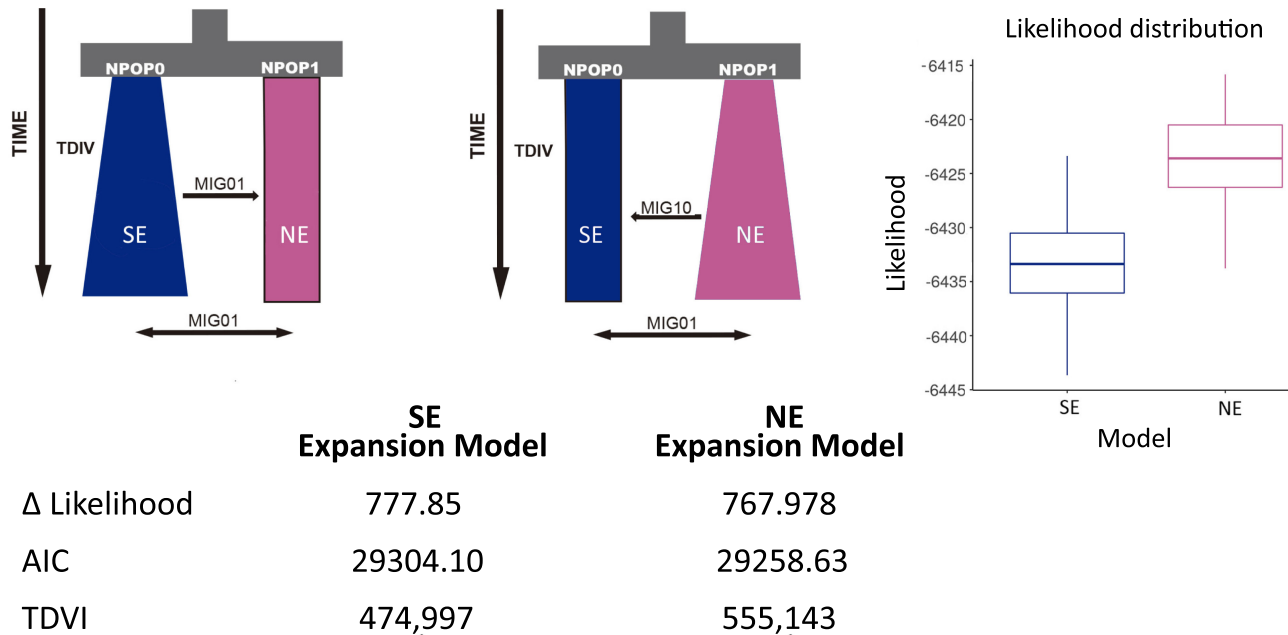


FIGURE 3 *Perinereis ponteni* demographic models simulated in Fastsimcoal (upper half) and their results (lower half) for both models assuming demographic expansion from the northeast region (NE, GBP+ItaP+PAP+SBP) with posterior migration to southeastern populations, and from SE region (AP+DuP+JP+PP) with posterior northward migration to the northeastern populations. Table shows the comparison between the results from each model. Boxplot presents the individual likelihoods of each model. The results from the northeastern expansion model also exhibited lower likelihood variation, providing more accuracy to the data. Each region is represented by the same colour in all graphics. TDVI, time of divergence between two populations estimated by Fastsimcoal. Araçá (AP), Praia dura (DuP), Jabaquara (JP), Prainha (PP), Itaoca (ItaP), Gamboa (GBP), Prado (PAP), Sabiaguaba (SBP).

and there was no overlap between RDA and BayeScan. All the selected variables from WorldClim were temperature-related and had 974 SNPs with frequencies associated with their variance. The variables from Bio-Oracle were nutrient-related primarily, except for the pH level at the sea surface, which had the second lowest number of associated SNPs. Only the RDA first axis was significant ($p = 0.05$), explaining 36% of genomic variation. According to the relative arrangement of the samples in the ordination space of axis 1, only the individual genotypes from the Praia Dura and Itaoca samples were positively correlated with high bathymetry. All other individual genotypes from the remaining localities were positively associated with lower bathymetry and higher annual mean temperature (Figure S7). The east-west and latitude aspects were also positively associated with a few genotypes from Prado and Gamboa. Most loci had no hits in the BLAST searches. However, we were able to identify genes responsible for intracellular transport (kinase heavy chain) and cytoskeletal dynamics (myosin), among others (Table S12).

Outlier SNPs were also used to investigate the population structure to evaluate any possible structuring related to environmental variables. The PCA results exhibited the same pattern for all loci (data not shown). Structure indicated the presence of three genetic clusters. The first comprised specimens from SE locations up to Rio de Janeiro, the second included specimens from the remaining localities and the third cluster was observed in all individuals (Figure S8). The Structure results using only putatively neutral SNPs revealed only two genetic clusters (Figure S9).

3.6 | Niche-based inferences

The simulations of the palaeodistribution of *P. ponteni* suggest the presence of an ecological refuge around Brazil's southeast ($\sim 22^{\circ}$ – 33° S) region during the LGM (~ 21 kya). This region, which includes the Brazilian coast from Rio de Janeiro to the Rio Grande do Sul, is the only one to present environmental conditions for species survival at that time (Figure 6a). The suitability distribution was drastically modified in the mid-Holocene (~ 6 kya), when the Southwest Atlantic coast from the Caribbean up to the north of Rio Grande do Sul (South region) was highly suitable for species survival (Figure 6b). The current spatial distribution simulation showed similar results, with all coasts being highly suitable for *P. ponteni* (Figure 6c). According to the suitability distribution, the southern coast of South America and the Amazonas River estuary show lower values of suitability, indicating that these areas might act as permeable barriers.

4 | DISCUSSION

We attempted to assess genomic diversity and possible factors influencing the population structure of a polychaete species, along with larval drift and niche modelling approaches. Our results indicate well-connected populations but an IBE pattern and less intense gene flow between the NE and SE regions. The findings also indicated a cline distribution pattern for the 14 putatively neutral SNPs in the

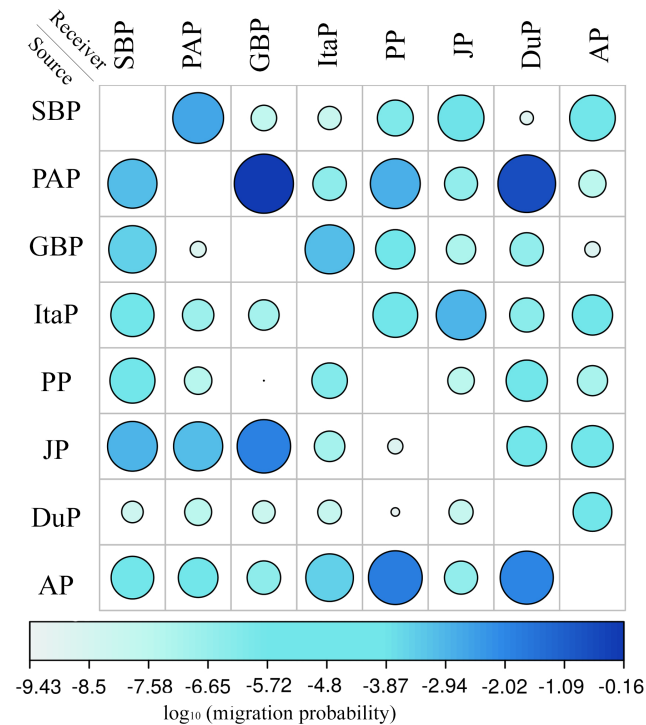


FIGURE 4 *Perinereis ponteni* migrant rates per generation from Fastsimcoal simulation with migration between all pairs of sites. Migration rate values are presented in the lower bar. The circle sizes and colours represent the rates of migrants from the source locality (rows) towards the receiver locality (columns). Araçá (AP), Praia dura (DuP), Jabaquara (JP), Prainha (PP), Itaoca (ItaP), Gamboa (GBP), Prado (PAP), Sabiaguaba (SBP).

southeastern region. Changes in allele distribution that characterize the cline occur in the same region where gene flow between populations is weaker. The palaeodistribution simulation results further suggest that this is the region where species distribution was restricted during the LGM.

According to Rouse (2000), all Nereididae species have lecithotrophic larvae with a shortened PLD until settlement compared to planktotrophic larvae (Hoegh-Guldberg & Emler, 1997). The PLD is long enough for larvae to travel thousands of kilometres, connecting the Brazilian populations carried by ocean currents. The detected low but significant value of the global F_{ST} (0.037) is an indication of limited connectivity. However, the absence of IBD and the presence of IBE (indicated by the significant correlation between genetic and temperature differences) suggest that the environment, and not the geographic distance, is responsible for such genetic differentiation. The presence of IBE in the absence of IBD was detected for other marine species, like the clownfish *Amphiprion bicinctus* (Nanninga et al., 2014; Saenz-Agudelo et al., 2015), the shore crab *Cyclograpsus punctatus* and the Cape urchin *Parechinus angulosus* (Nielsen et al., 2020). The difference in temperature in this case seems to affect not only the dispersion phase (i.e. larvae) but mainly the adult phase when *P. ponteni* has lower motility and is exposed to local conditions for extended periods (Bakken et al., 2018). In addition, the temperature can affect these animals directly and indirectly by changing the amount

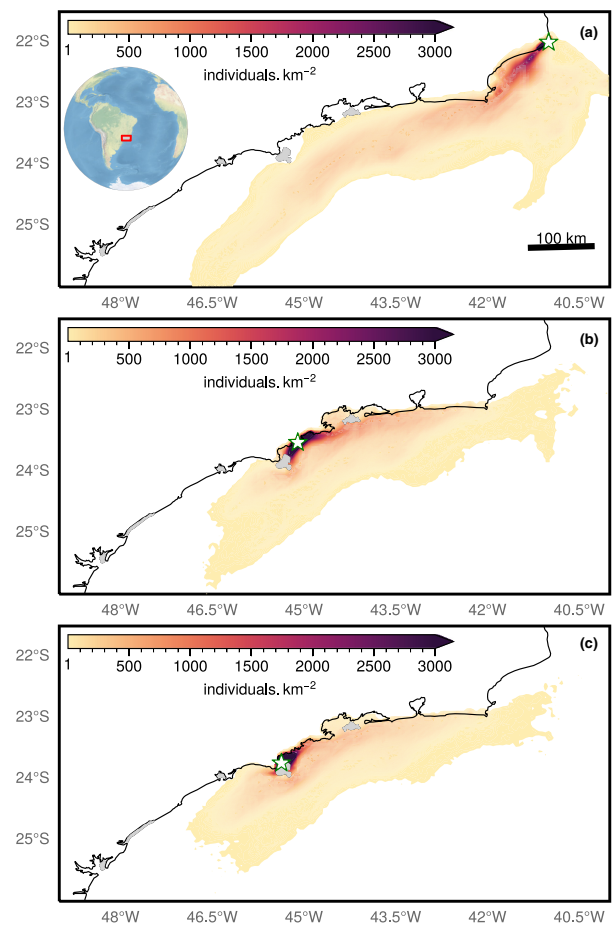


FIGURE 5 Population density considering 1000 daily seeded individuals in the vicinity of (a) Itaoca—ES, (b) Praia Dura—SP and (c) Araçá—SP. A hydrodynamic simulation PreAMar (Costa et al., 2020) provides the currents, salinity and temperature for the year of 2019. Each individual is viable for an interval of 50 days. The white-green stars represent the *Perinereis ponteni* seeding/sampling sites. The projection of the main map is the Equirectangular.

of dissolved O_2 and microbiota abundance (White et al., 1991). Structural results support this possibility, highlighting a boundary for migration in the northern limit of the SE region around Cabo São Tomé, where something seems to act as a boundary for connectivity. Structure, PCA, cline analyses and oceanic circulation models indicated that this boundary is located around the Cabo Frio region.

The Cabo Frio region, which is the most known region of the upwelling system, seems to be a gene flow barrier or a sieve for many marine species, such as the crustaceans *Excirrolana braziliensis* and *Litopenaeus schmitti* (Hurtado et al., 2016; Maggioni et al., 2003), the coral *Mussismilia hispida* (Peluso et al., 2018), the fish *Mugil liza* (Mai et al., 2014) and even the dolphin *Pontoporia blainvillei* (Lázaro et al., 2004). Mainly during the summer, this region receives strong winds from the continental area that push the warm coastal waters, giving rise to deeper cold waters rich in nutrients and forming a constant upwelling (Martins et al., 2021; Valentin et al., 1987). Therefore, the temperature and nutrient regimes in this area are quite different from those in subjacent zones and are thought to alter the local

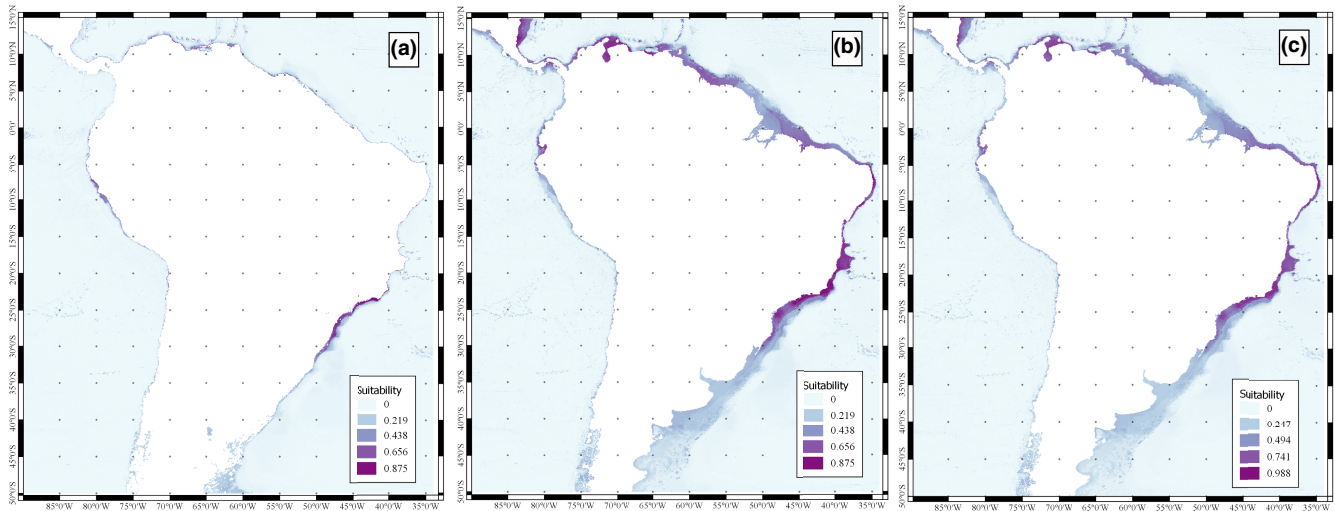


FIGURE 6 Niche modelling for palaeodistribution of *Perinereis ponteni* from (a) last glacial maximum, ~21 kya; (b) mid-Holocene, ~6 kya; and (c) present time, 0 kya. Different colours represent different values of suitability for the species as depicted in each image. The projection is transverse Mercator.

community and act as a semipermeable boundary for some species (Mai et al., 2014; Martins et al., 2021; Peluso et al., 2018; Volk et al., 2021). In addition, our results indicate a hydrodynamic barrier preventing larval exchange between adjacent areas (such as Prainha and Gamboa) rather than an ecological barrier. Hydrodynamic barriers are of cause isolation among coral species between the Great Barrier Reef and Lord Howe Island (Keith et al., 2015; Wood et al., 2014), which might also be the cause of the pattern seen here. The larval dispersion simulations showed virtually no dispersion across the Cabo Frio region, but our demographic results showed some migration among these sites. This sign could be an effect of past migration or climatic anomalies. The populations of *P. ponteni* from Espírito Santo (Gamboa and Itaoca) and Rio de Janeiro (Prainha and Jabaquara) seemed to have a stronger response to the Cabo Frio barrier, presenting significant pairwise values of F_{ST} , despite their proximity. However, this population structure was not detected in a previous phylogeographic study using mitochondrial markers from Paiva et al. (2019). The authors analysed the population structure of *P. ponteni* along the Brazilian coast, and the populations presented panmixia with a few shared haplotypes distributed along the entire coast. Such disparity between genomic and mitochondrial data can be an effect of sex-biased inheritance associated with the different rates of mutation in different markers (Pazmiño et al., 2017; Toews & Brelsford, 2012), but most likely are a result of the large number of markers analysed in this study.

Gene flow along the Brazilian coast, resulting in successive colonization events from different cohorts and generations, might end in mixed genetically distinct subpopulations along the species distribution. This would explain the heterozygosity deficiency found with statistically significant values of F_{IS} . Heterozygote deficiencies are commonly found in marine invertebrates and can result from high rates of inbreeding in small populations, natural selection or the Wahlund effect (Castric et al., 2002; Whitaker, 2004).

4.1 | Palaeodistribution and origin of present populations

A previous study on mitochondrial genes found evidence of an important founder effect in *P. ponteni* populations along the Brazilian coast (Paiva et al., 2019). This is congruent with our findings in palaeodistribution simulations. These simulations suggest that the environmental conditions necessary for *P. ponteni* survival were restricted to the southeastern coast of Brazil during the LGM (~21 kya), which acts as a refuge for the species. However, the demographic simulations in Fastsimcoal show that the northern coastal region is the most probable dispersion centre for the species. Since the generation time of Nereididae varies considerably, from a couple of months to years, and it is directly affected by environmental temperature (Hardege, 1999), we cannot pinpoint the time of divergence between populations in years. However, palaeodistribution simulations suggest that *P. ponteni* distribution was restricted to the southern coast of Brazil during the LGM, expanding northward and southward during the mid-Holocene. The same pattern was found for other coastal species in Brazil (Ayres-Ostrock et al., 2019; Cortez et al., 2021), and given the significant change in oceanic temperature and sea level from the LGM to the present (Tierney et al., 2020), such restriction in the distribution is expected. The climatic conditions in the mid-Holocene along the range of *P. ponteni* distribution are akin to the present conditions (Steig, 1999). The ocean currents and ecological factors were also similar to the present ones (Gu et al., 2018), and the species distribution in this period was very similar to the present.

4.2 | Signals of adaptive selection

Of the 23,300 SNPs analysed, 1571 were supposedly under selection according to BayeScan, RDA and LEA, representing over 450

loci (Table S13). RDA and LFMM have a similar power to identify genetic markers related to environmental variables (Capblancq et al., 2018). RDA has the advantage of being able to detect the main environmental gradients. In addition to IBE, indicating the importance of temperature in these populations, a potential association between allele frequency and temperature was supported by the RDA and LFMM, indicating that this variable also has a high contribution to the genotype distribution, according to the pattern observed in Figure S7. Temperature variation was already reported as being capable of driving population divergences for invertebrate species across different geographic scales (Benestan et al., 2016; Miller et al., 2019; Takeuchi et al., 2020). Moreover, other predictors seem to be relevant to the distribution of allele frequencies, such as salinity, bathymetry and curvature profile. Such variables play an essential role in the performance and survival of species with planktonic larval stages, especially for those who inhabit rocky shores. Furthermore, three-dimensional features, such as curvature profile, can cause widespread impacts on biodiversity and functional traits (Borland et al., 2022; Demopoulos et al., 2018), since it influences water flow dynamics across the surface. These assumptions are related to how geomorphic metrics can distinguish different benthic communities, even for polychaetes (De Leo et al., 2014), showing again that environmental heterogeneity, rather than only the geographic distance, is responsible for the genetic differentiation (Harley & Helmuth, 2003). Local adaptation is a common pattern among marine invertebrates with low motility in the adult phase because of the strong influence of local conditions (de Wit & Palumbi, 2013; Leiva et al., 2019; Sandoval-Castillo et al., 2017). Our inability to annotate many outlier loci was probably due to the lack of annotated Nereididae genomes. Changes in the expression of myosin and kinase genes were detected in zebrafish and the larvae of *Sparus aurata* (Chen et al., 2013; Gavrilova et al., 2013; Madeira et al., 2017; Tomanek, 2014), in response to hypoxia and heat shock. The signs of adaptation involving these genes further indicate the importance of temperature in *P. ponteni* individuals.

This study sheds light on the influence of temperature on connectivity among Brazilian populations of *P. ponteni*, as indicated by the genetic boundary in the upwelling region. Many outlier loci discovered here were directly (971 SNPs) or indirectly (863 SNPs) related to temperature. This is not only due to the disruption of local conditions but also because ocean currents are determined by the differences in ocean temperature and salinity, both of which can change dramatically in the coming years (Voosen, 2020). This change in the velocity and direction of the currents will most likely affect the connection between populations, affecting gene flow in the species. Not only does temperature affect species survival or fitness, but other environmental factors such as primary production may also significantly affect *Perinereis* fitness. This study reveals that environmental changes can also affect polychaetes at the genomic level. It might be an interesting model for analysing the oceanic warming effect and its consequences on natural populations.

ACKNOWLEDGEMENTS

This study was financed by São Paulo Research Foundation (FAPESP) grants 2015/20139-9 and 2016/20005-5, and by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) Finance Code 001. Specimens were collected under permits issued by the Institute Chico Mendes (ICMBio) (numbers 55701 and 67004). The authors are grateful to Paulo Pachele P. Gurgel, Felipe A.C. Monteiro, Tuane Ribeiro, Stephanie Pruffer, Gabriel Sonoda and Helena Mathews Cascon for helping with field trips and material collection; Priscilla Villela and Ecomol Consultoria e Projetos for help with library construction and genomic sequencing; Horácio Montenegro for helping with analyses; Gabriel Marroig, Diogo Melo and Vitor Aguiar for helping with the use of the Darwin server; CEBIMar-USP (Centro de Biologia Marinha, Universidade de São Paulo) and its staff for providing the essential laboratory facilities and logistics. We would like to thank Editage (www.editage.com) for the English language editing and both reviewers for their insightful comments, which improved the manuscript.

CONFLICT OF INTEREST

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as potential conflicts of interest.

DATA AVAILABILITY STATEMENT

The datasets analysed in this study can be found in the NCBI SRA database (<https://www.ncbi.nlm.nih.gov/sra>) under the PRJNA750868 code. Complete set of data input: Dryad <https://doi.org/10.5061/dryad.8gtht76s9>. New distribution *Perinereis ponteni* data deposited at GBIF: <https://doi.org/10.15468/dl.e2ken2>.

ORCID

Cecili B. Mendes  <https://orcid.org/0000-0002-9041-1270>

Thainá Cortez  <https://orcid.org/0000-0003-2312-5087>

Cinthya S. G. Santos  <https://orcid.org/0000-0003-2182-0801>

Thadeu Sobral-Souza  <https://orcid.org/0000-0002-2367-2315>

Arian Dielactaquiz Santos  <https://orcid.org/0000-0002-7746-8205>

Dalton Kei Sasaki  <https://orcid.org/0000-0002-8006-0200>

Danilo Augusto Silva  <https://orcid.org/0000-0002-9714-4937>

Marcelo Dottori  <https://orcid.org/0000-0003-2382-4136>

Sônia C. S. Andrade  <https://orcid.org/0000-0002-1302-5261>

REFERENCES

- Allouche, O., Tsoar, A., & Kadmon, R. (2006). Assessing the accuracy of species distribution models: Prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology*, 43(6), 1223–1232.
- Alvarado-Serrano, D. F., & Knowles, L. L. (2014). Ecological niche models in phylogeographic studies: Applications, advances and precautions. *Molecular Ecology Resources*, 14(2), 233–248.
- Araújo, M. B., Anderson, R. P., Barbosa, A. M., Beale, C. M., Dormann, C. F., Early, R., Garcia, R. A., Guisan, A., Maiorano, L., Naimi, B., O'Hara, R. B., Zimmermann, N. E., & Rahbek, C. (2019). Standards

- for distribution models in biodiversity assessments. *Science Advances*, 5(1), eaat4858.
- Ayres-Ostrock, L. M., Valero, M., Mauger, S., Oliveira, M. C., Plastino, E. M., Guillemin, M. L., & Destombe, C. (2019). Dual influence of terrestrial and marine historical processes on the phylogeography of the Brazilian intertidal red alga *Gracilaria caudata*. *Journal of Phycology*, 55(5), 1096–1114.
- Bakken, T., Glasby, C. J., Santos, C. S. G., & Wilson, R. S. R. S. (2018). Nereididae Blainville, 1818. In W. Westheide, G. Purschke, & M. Böggemann (Eds.), *Handbook of zoology - a natural history of the phyla of the animal kingdom - Annelida: Polychaetes* (pp. 1–43). De Gruyter.
- Benestan, L., Quinn, B. K., Maaroufi, H., Laporte, M., Clark, F. K., Greenwood, S. J., Rochette, R., & Bernatchez, L. (2016). Seascape genomics provides evidence for thermal adaptation and current-mediated population structure in American lobster (*Homarus americanus*). *Molecular Ecology*, 25(20), 5073–5092.
- Bernardi, G., Sordino, P., & Powers, D. A. (1993). Concordant mitochondrial and nuclear DNA phylogenies for populations of the teleost fish *Fundulus heteroclitus*. *Proceedings of the National Academy of Sciences*, 90(20), 9271–9274.
- Bernatchez, S., Xuereb, A., Laporte, M., Benestan, L., Steeves, R., Laflamme, M., Bernatchez, S., & Mallet, M. A. (2018). Seascape genomics of eastern oyster (*Crassostrea virginica*) along the Atlantic coast of Canada. *Evolutionary Applications*, 12(3), 587–609.
- Blumberg, A., Georgas, N., Herrington, T., & Bruno, M. (2010). How good is good? Establishing performance goals for marine forecast systems. *Ninth conference on coastal atmospheric and oceanic prediction and processes*, Annapolis, MD.
- Bohonak, A. J. (1999). Dispersal, gene flow, and population structure. *Quarterly Review of Biology*, 74(1), 21–45.
- Borland, H. P., Gilby, B. L., Henderson, C. J., Connolly, R. M., Gorissen, B., Ortodossi, N. L., Rummell, A. J., Pittman, S. J., Sheaves, M., & Olds, A. D. (2022). Dredging fundamentally reshapes the ecological significance of 3D terrain features for fish in estuarine seascapes. *Landscape Ecology*, 37, 1385–1400.
- Caley, M. J., Carr, M. H., Hixon, M. A., Hughes, T. P., Jones, G. P., & Menge, B. A. (1996). Recruitment and the local dynamics of open marine populations. *Annual Review of Ecology and Systematics*, 27(1), 477–500.
- Capblancq, T., Luu, K., Blum, M. G., & Bazin, E. (2018). Evaluation of redundancy analysis to identify signatures of local adaptation. *Molecular Ecology Resources*, 18(6), 1223–1233.
- Castric, V., Bernatchez, L., Belkhir, K., & Bonhomme, F. (2002). Heterozygote deficiencies in small lacustrine populations of brook charr *Salvelinus fontinalis* Mitchell (Pisces, Salmonidae): A test of alternative hypotheses. *Heredity*, 89(1), 27–35.
- Chen, K., Cole, R. B., & Rees, B. B. (2013). Hypoxia-induced changes in the zebrafish (*Danio rerio*) skeletal muscle proteome. *Journal of Proteomics*, 78, 477–485.
- Cortez, T., Amaral, R. V., Sobral-Souza, T., & Andrade, S. (2021). Genome-wide assessment elucidates connectivity and the evolutionary history of the highly dispersive marine invertebrate *Littoraria flava* (Littorinidae: Gastropoda). *Biological Journal of the Linnean Society*, blab055.
- Costa, C. G. R., Leite, J. R. B., Castro, B. M., Blumberg, A. F., Georgas, N., Dottori, M., & Jordi, A. (2020). An operational forecasting system for physical processes in the Santos-São Vicente-Bertioga estuarine system, Southeast Brazil. *Ocean Dynamics*, 70(2), 257–271.
- Dagestad, K. F., Röhrs, J., Breivik, Ø., & Ådlandsvik, B. (2017). OpenDrift v1.0: A generic framework for trajectory modeling. *Geoscientific Model Development Discussion*, 11, 1–28. <https://doi.org/10.5194/gmd-2017-205>
- Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., Durbin, R., & 1000 Genomes Project Analysis Group. (2011). The variant call format and VCFtools. *Bioinformatics*, 27(15), 2156–2158.
- De Leo, F. C., Vetter, E. W., Smith, C. R., Rowden, A. A., & McGranaghan, M. (2014). Spatial scale-dependent habitat heterogeneity influences submarine canyon macrofaunal abundance and diversity off the Main and northwest Hawaiian islands. *Deep Sea Research Part II: Topical Studies in Oceanography*, 104, 267–290.
- de Medeiros, B. A., & Farrell, B. D. (2018). Whole-genome amplification in double-digest RADseq results in adequate libraries but fewer sequenced loci. *PeerJ*, 6, e5089.
- de Wit, P., & Palumbi, S. R. (2013). Transcriptome-wide polymorphisms of red abalone (*Haliotis rufescens*) reveal patterns of gene flow and local adaptation. *Molecular Ecology*, 22(11), 2884–2897.
- Demopoulos, A. W., Bourque, J. R., Durkin, A., & Cordes, E. E. (2018). The influence of seep habitats on sediment macrofaunal biodiversity and functional traits. *Deep Sea Research Part I: Oceanographic Research Papers*, 142, 77–93.
- Derryberry, E. P., Derryberry, G. E., Maley, J. M., & Brumfield, R. T. (2014). HZAR: Hybrid zone analysis using an R software package. *Molecular Ecology Resources*, 14(3), 652–663.
- Díaz-Castañeda, V., & Reish, D. (2009). *Polychaetes in environmental studies. Annelids as model systems in the biological sciences*. John Wiley & Sons.
- Doyle, J. J., & Doyle, J. L. (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin*, 19(1), 11–15.
- Earl, D. A. (2012). STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, 4(2), 359–361.
- Eaton, D. A., & Overcast, I. (2020). Ipyrad: Interactive assembly and analysis of RADseq datasets. *Bioinformatics*, 36(8), 2592–2594.
- Edwards, K. P., Werner, F. E., & Blanton, B. O. (2006). Comparison of observed and modeled drifter trajectories in coastal regions: An improvement through adjustments for observed drifter slip and errors in wind fields. *Journal of Atmospheric and Oceanic Technology*, 23(11), 1614–1620.
- Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., & Mitchell, S. E. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS One*, 6(5), e19379.
- Excoffier, L., & Foll, M. (2011). Fastsimcoal: A continuous-time coalescent simulator of genomic diversity under arbitrarily complex evolutionary scenarios. *Bioinformatics*, 27(9), 1332–1334.
- Excoffier, L., & Lischer, H. E. (2010). Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and windows. *Molecular Ecology Resources*, 10(3), 564–567.
- Fauchald, K., & Jumars, P. A. (1979). The diet of worms: A study of polychaete feeding guilds. *Oceanography and Marine Biology Annual Review*, 17, 193–284.
- Ferrier, S., & Guisan, A. (2006). Spatial modelling of biodiversity at the community level. *Journal of Applied Ecology*, 43(3), 393–404.
- Folkvord, A. (2005). Comparison of size-at-age of larval Atlantic cod (*Gadus Morhua*) from different populations based on size- and temperature-dependent growth models. *Canadian Journal of Fisheries and Aquatic Sciences*, 62(5), 1037–1052.
- 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(2), 977–993.
- Forester, B. R., Lasky, J. R., Wagner, H. H., & Urban, D. L. (2018). Comparing methods for detecting multilocus adaptation with multivariate genotype–environment associations. *Molecular Ecology*, 27(9), 2215–2233.
- Francis, R. M. (2017). Pophelper: A R package and web app to analyse and visualize population structure. *Molecular Ecology Resources*, 17(1), 27–32.
- Frichot, E., & François, O. (2015). LEA: An R package for landscape and ecological association studies. *Methods in Ecology and Evolution*, 6(8), 925–929.

- Frichot, E., Schoville, S. D., Bouchard, G., & François, O. (2013). Testing for associations between loci and environmental gradients using latent factor mixed models. *Molecular Biology and Evolution*, 30(7), 1687–1699.
- Gavrilova, L. P., Korpacheva, I. I., Semushina, S. G., & Yashin, V. A. (2013). Heat shock induces simultaneous rearrangements of all known cytoskeletal filaments in normal interphase fibroblasts. *Cell and Tissue Biology*, 7(1), 54–63.
- Gu, F., Chiessi, C. M., Zonneveld, K. A., & Behling, H. (2018). Late quaternary environmental dynamics inferred from marine sediment core GeoB6211-2 off southern Brazil. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 496, 48–61.
- Hardege, J. D. (1999). Nereidid polychaetes as model organisms for marine chemical ecology. *Hydrobiologia*, 402, 145–161.
- Harley, C. D. G., & Helmuth, B. S. T. (2003). Local- and regional-scale effects of wave exposure, thermal stress, and absolute versus effective shore level on patterns of intertidal zonation. *Limnology and Oceanography*, 48, 1498–1508.
- Hoegh-Guldberg, O., & Emler, R. B. (1997). Energy use during the development of a lecithotrophic and a planktotrophic echinoid. *The Biological Bulletin*, 192(1), 27–40.
- Hurtado, L. A., Mateos, M., Mattos, G., Liu, S., Haye, P. A., & Paiva, P. C. (2016). Multiple transisthmian divergences, extensive cryptic diversity, occasional long-distance dispersal, and biogeographic patterns in a marine coastal isopod with an ampho-American distribution. *Ecology and Evolution*, 6(21), 7794–7808.
- Jombart, T. (2008). ADEGENET: A R package for the multivariate analysis of genetic markers. *Bioinformatics*, 24(11), 1403–1405.
- Keith, S. A., Woolsey, E. S., Madin, J. S., Byrne, M., & Baird, A. H. (2015). Differential establishment potential of species predicts a shift in coral assemblage structure across a biogeographic barrier. *Ecography*, 38(12), 1225–1234.
- Lázaro, M., Lessa, E. P., & Hamilton, H. (2004). Geographic genetic structure in the franciscana dolphin (*Pontoporia blainvillei*). *Marine Mammal Science*, 20(2), 201–214.
- Leiva, C., Taboada, S., Kenny, N. J., Combosch, D., Giribet, G., Jombart, T., & Riesgo, A. (2019). Population substructure and signals of divergent adaptive selection despite admixture in the sponge *Dendrilla antarctica* from shallow waters surrounding the Antarctic peninsula. *Molecular Ecology*, 28(13), 3151–3170.
- Levin, L. A. (2006). Recent progress in understanding larval dispersal: New directions and digressions. *Integrative and Comparative Biology*, 46(3), 282–297.
- Lischer, H. E., & Excoffier, L. (2012). PGDSpider: An automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics*, 28(2), 298–299.
- Madeira, D., Araujo, J. E., Vitorino, R., Costa, P. M., Capelo, J. L., Vinagre, C., & Diniz, M. S. (2017). Molecular plasticity under ocean warming: Proteomics and fitness data provides clues for a better understanding of the thermal tolerance in fish. *Frontiers in Physiology*, 8, 825.
- Maggioni, R., Rogers, A. D., & Maclean, N. (2003). Population structure of *Litopenaeus schmitti* (Decapoda: Penaeidae) from the Brazilian coast identified using six polymorphic microsatellite loci. *Molecular Ecology*, 12(12), 3213–3217.
- Mai, A. C., Mino, C. I., Marins, L. F., Monteiro-Neto, C., Miranda, L., Schwingel, P. R., Lemos, V. M., Gonzalez-Castro, M., Castello, J. P., & Vieira, J. P. (2014). Microsatellite variation and genetic structuring in *Mugil liza* (Teleostei: Mugilidae) populations from Argentina and Brazil. *Estuarine, Coastal and Shelf Science*, 149, 80–86.
- Mantel, N. (1967). The detection of disease clustering and a generalized regression approach. *Cancer Research*, 27(2 Part 1), 209–220.
- Martins, N., Macagnan, L., Cassano, V., & Gurgel, C. (2021). Barriers to gene flow along the Brazilian coast: A synthesis and data analysis. *Authorea* [Preprints]. <https://doi.org/10.22541/au.161358135.51187023/v1>
- Miller, A. D., Hoffmann, A. A., Tan, M. H., Young, M., Ahrens, C., Cocomazzo, M., Rattray, A., Ierodiaconou, D. A., Trembl, E., & Sherman, C. D. (2019). Local and regional scale habitat heterogeneity contribute to genetic adaptation in a commercially important marine mollusc (*Haliotis rubra*) from southeastern Australia. *Molecular Ecology*, 28(12), 3053–3072.
- Nanninga, G. B., Saenz-Agudelo, P., Manica, A., & Berumen, M. L. (2014). Environmental gradients predict the genetic population structure of a coral reef fish in the Red Sea. *Molecular Ecology*, 23(3), 591–602.
- Nielsen, E. S., Henriques, R., Beger, M., Toonen, R. J., & Von der Heyden, S. (2020). Multi-model seascape genomics identifies distinct environmental drivers of selection among sympatric marine species. *BMC Evolutionary Biology*, 20(1), 1–17.
- Nunes, F. L., Rigal, F., Dubois, S. F., & Viard, F. (2021). Looking for diversity in all the right places? Genetic diversity is highest in peripheral populations of the reef-building polychaete *Sabellaria alveolata*. *Marine Biology*, 168(5), 1–16.
- Nunes, J. D. R. S., Liu, S., Pértille, F., Perazza, C. A., Villela, P. M. S., de Almeida-Val, V. M. F., Hilsdorf, A. W. S., Liu, Z., & Coutinho, L. L. (2017). Large-scale SNP discovery and construction of a high-density genetic map of *Colossoma macropomum* through genotyping-by-sequencing. *Scientific Reports*, 7(1), 1–11.
- Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., O'hara, R. B., Solymos, P., Stevens, M. H. H., Szoecs, E., Wagner, H., Barbour, M., Bedward, M., Bolker, B., Borcard, D., Carvalho, G., Chirico, M., De Caceres, M., Durand, S., ... Wagner, H. (2020). Vegan: community ecology package. R package version 2.3-0. *Scientific Reports*, 10, 20354.
- Paiva, P. C., Mutaquilha, B. F., Coutinho, M. C. L., & Santos, C. S. (2019). Comparative phylogeography of two coastal species of *Perinereis* Kinberg, 1865 (Annelida, Polychaeta) in the South Atlantic. *Marine Biodiversity*, 49(3), 1537–1551.
- Palumbi, S. R. (1994). Genetic divergence, reproductive isolation, and marine speciation. *Annual Review of Ecology and Systematics*, 25(1), 547–572.
- Pamungkas, J., Glasby, C. J., Read, G. B., Wilson, S. P., & Costello, M. J. (2019). Progress and perspectives in the discovery of polychaete worms (Annelida) of the world. *Helgoland Marine Research*, 73(1), 1–10.
- 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(5), e0176419.
- Pazmiño, D. A., Maes, G. E., Simpfendorfer, C. A., Salinas-de-León, P., & van Herwerden, L. (2017). Genome-wide SNPs reveal low effective population size within confined management units of the highly vagile Galapagos shark (*Carcharhinus galapagensis*). *Conservation Genetics*, 18(5), 1151–1163.
- Peluso, L., Tascheri, V., Nunes, F. L., Castro, C. B., Pires, D. O., & Zilberberg, C. (2018). Contemporary and historical oceanographic processes explain genetic connectivity in a southwestern Atlantic coral. *Scientific Reports*, 8(1), 1–12.
- Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, 155(2), 945–959.
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., Maller, J., Sklar, P., de Bakker, P. I., Daly, M. J., & Sham, P. C. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *The American Journal of Human Genetics*, 81(3), 559–575.
- 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(17), 4348–4370.
- Rouse, G., & Pleijel, F. (2001). *Polychaetes*. Oxford University Press.
- Rouse, G. W. (2000). Polychaetes have evolved feeding larvae numerous times. *Bulletin of Marine Science*, 67(1), 391–409.

- Saenz-Agudelo, P., Dibattista, J. D., Piatek, M. J., Gaither, M. R., Harrison, H. B., Nanninga, G. B., & Berumen, M. L. (2015). Seascape genetics along environmental gradients in the Arabian peninsula: Insights from ddRAD sequencing of anemonefishes. *Molecular Ecology*, 24(24), 6241–6255.
- Sandoval-Castillo, J., Robinson, N. A., Hart, A. M., Strain, L. W., & Beheregaray, L. B. (2017). Seascape genomics reveals adaptive divergence in a connected and commercially important mollusc, the greenlip abalone (*Haliotis laevis*), along a longitudinal environmental gradient. *Molecular Ecology*, 27(7), 1603–1620.
- Sbrocco, E. J. (2014). Paleo-MARSPEC: Gridded ocean climate layers for the mid-Holocene and last glacial maximum: Ecological archives E095-149. *Ecology*, 95(6), 1710.
- Sbrocco, E. J., & Barber, P. H. (2013). MARSPEC: Ocean climate layers for marine spatial ecology: Ecological archives E094-086. *Ecology*, 94(4), 979.
- Selkoe, K. A., Aloia, C. C., Crandall, E. D., Iacchi, M., Liggins, L., Puritz, J. B., von der Hyden, S., & Toonen, R. J. (2016). A decade of seascape genetics: Contributions to basic and applied marine connectivity. *Marine Ecology Progress Series*, 554, 1–19.
- Steig, E. J. (1999). Mid-Holocene climate change. *Science*, 286(5444), 1485–1487.
- Stigall, A. L. (2012). Using ecological niche modeling to evaluate niche stability in deep time. *Journal of Biogeography*, 39(4), 772–781.
- Sundby, S. (1983). A one-dimensional model for the vertical distribution of pelagic fish eggs in the mixed layer. *Deep Sea Research Part A. Oceanographic Research Papers*, 30(6), 645–661.
- Storfer, A., Patton, A., & Fraik, A. K. (2018). Navigating the interface between landscape genetics and landscape genomics. *Frontiers in Genetics*, 9(68), 1–14. <https://doi.org/10.3389/fgene.2018.00068>
- Takeuchi, T., Masaoka, T., Aoki, H., Koyanagi, R., Fujie, M., & Satoh, N. (2020). Divergent northern and southern populations and demographic history of the pearl oyster in the western Pacific revealed with genomic SNPs. *Evolutionary Applications*, 13(4), 837–853.
- Tierney, J. E., Zhu, J., King, J., Malevich, S. B., Hakim, G. J., & Poulsen, C. J. (2020). Glacial cooling and climate sensitivity revisited. *Nature*, 584(7822), 569–573.
- Toews, D. P., & Brelsford, A. (2012). The biogeography of mitochondrial and nuclear discordance in animals. *Molecular Ecology*, 21(16), 3907–3930.
- Tomanek, L. (2014). Proteomics to study adaptations in marine organisms to environmental stress. *Journal of Proteomics*, 105, 92–106.
- Valentin, J. L., Andre, D. L., & Jacob, S. A. (1987). Hydrobiology in the Cabo Frio (Brazil) upwelling: Two-dimensional structure and variability during a wind cycle. *Continental Shelf Research*, 7(1), 77–88.
- Volk, D. R., Konvalina, J. D., Floeter, S. R., Ferreira, C. E., & Hoffman, E. A. (2021). Going against the flow: Barriers to gene flow impact patterns of connectivity in cryptic coral reef gobies throughout the western Atlantic. *Journal of Biogeography*, 48(2), 427–439.
- Voosen, P. (2020). Climate change spurs global speedup of ocean currents. *Science*, 367(6478), 612–613.
- Wang, I. J., & Bradburd, G. S. (2014). Isolation by environment. *Molecular Ecology*, 23(23), 5649–5662.
- Weir, B. S., & Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution*, 38(6), 1358–1370.
- Whitaker, K. (2004). Non-random mating and population genetic subdivision of two broadcasting corals at Ningaloo reef, Western Australia. *Marine Biology*, 144(3), 593–603.
- White, P. A., Kalff, J., Rasmussen, J. B., & Gasol, J. M. (1991). The effect of temperature and algal biomass on bacterial production and specific growth rate in freshwater and marine habitats. *Microbial Ecology*, 21(1), 99–118.
- Wood, S., Paris, C. B., Ridgwell, A., & Hendy, E. J. (2014). Modelling dispersal and connectivity of broadcast spawning corals at the global scale. *Global Ecology and Biogeography*, 23(1), 1–11.
- Zhbannikov, I. Y., Hunter, S. S., Foster, J. A., & Settles, M. L. (2017). SeqClean: A pipeline for high-throughput sequence data pre-processing. In *Proceedings of the 8th ACM international conference on bioinformatics, computational Biology, and health informatics* (pp. 407–416). ACM.

BIOSKETCHES

Cecili B. Mendes is interested in the population genomics of marine worms, especially polychaetes and ribbon worms, to investigate the environmental and developmental causes of the present patterns. She also works with nemertean taxonomy and diversity.

Thainá Cortez is interested in the population genomics of marine invertebrates, especially meiofauna or other microscopic lineages. She also works with local adaptation and selective forces investigated with molecular and bioinformatics tools to understand the environmental influence over the genomic diversity.

Author contributions: Cecili B. Mendes collected most specimens, performed lab work and genomic analyses wrote the first draft. Thainá Cortez performed genomic analyses, wrote the revision and revised the manuscript. Cinthya S. G. Santos wrote and revised the manuscript accordingly. Thadeu Sobral-Souza performed niche-based models and palaeodistribution scenario modelling and revised the manuscript. Arian Dielactaquiz Santos, Dalton Kei Sasaki, Danilo Augusto Silva and Marcelo Dottori performed larval dispersion simulations and revised the manuscript. Sônia C. S. Andrade collected specimens, supervised the genomic analyses and wrote and revised the manuscript. Sônia C. S. Andrade and Cecili B. Mendes conceived the project. All authors contributed to the manuscript writing and approved the submitted version.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Mendes, C. B., Cortez, T., Santos, C. S. G., Sobral-Souza, T., Santos, A. D., Sasaki, D. K., Silva, D. A., Dottori, M., & Andrade, S. C. S. (2022). Seascape genetics in a polychaete worm: Disentangling the roles of a biogeographic barrier and environmental factors. *Journal of Biogeography*, 00, 1–13. <https://doi.org/10.1111/jbi.14504>