



OPEN

Spatial and temporal stability in the genetic structure of a marine crab despite a biogeographic break

David Veliz^{1,2}✉, Noemi Rojas-Hernández¹, Caren Vega-Retter¹, Camila Zaviezo¹, Ignacio Garrido^{3,4,5} & Luis Miguel Pardo^{3,4}

Elucidating the processes responsible for maintaining the population connectivity of marine benthic species mediated by larval dispersal remains a fundamental question in marine ecology and fishery management. Understanding these processes becomes particularly important in areas with a biogeographic break and unidirectional water movement along the sides of the break. Based on variability at 4209 single-nucleotide polymorphisms in 234 individuals, we determine the genetic structure, temporal genetic stability, and gene flow among populations of the commercially important mola rock crab *Metacarcinus edwardsii* in a system in southern Chile with a biogeographic break at latitude 42°S. Specimens were collected at eight sites within its geographic distribution, with collection at four of these sites was performed twice. Using population genetic approaches, we found no evidence of geographic or temporal population differentiation. Similarly, we found no evidence of an effect on gene flow of the biogeographic break caused by the the West Wind Drift Current. Moreover, migration analyses supported gene flow among all sites but at different rates for different pairs of sites. Overall, our findings indicate that *M. edwardsii* comprises a single large population with high levels of gene flow among sites separated by over 1700 km and demonstrate temporal stability in its genetic structure.

Most adult benthic marine invertebrates have limited mobility or are sessile, as the larval stage is the period that enables dispersion¹. Although the larval period only represents approximately 1–5% of the total life cycle of these species, it is important for population connectivity² because it is fundamental to maintaining the cohesiveness of benthic populations, enabling them to persist through ecological and evolutionary time³.

Factors influencing marine population connectivity and larval dispersal include planktonic larval duration, larval behavior, and oceanographic currents^{4,5}. Planktonic larval duration is positively correlated with dispersal distance⁶ and depends on the species, ranging from merely hours among corals⁷ to a year among lobsters⁸. In addition, larval behavior contributes to changes in magnitude and direction of larval movement, enabling offshore transport or residence near the coast⁹. One of the most important larval behaviors described in the water column is active vertical migration¹⁰. This behavior could regulate larval engagement with physical forcing and circulation, which would prevent them from straying from the coast¹¹. Finally, all scales of oceanic water movement affect larval dispersal. Small-scale processes (e.g., turbulence, small eddies, and stagnant zones) combine with nearshore physical processes (e.g., waves, winds, and tides) to affect the degree of larval retention in each system¹², while mesoscale behaviors (e.g., meanders) and global-scale physical processes (e.g., the main oceanographic currents) affect the degree of larval dispersal and population connectivity at a large scale^{5,13}.

At large scales, the hydrographic regime and bottom topography affect biogeographic breaks. Examples include Cape Cod in the Atlantic Ocean¹⁴, Point Conception in North America¹⁵, and latitudes 30°S and 42°S in Chile¹⁶. The biogeographic break at 30°S is likely caused by differences in eddy kinetic energy (high south of and low north of 30°S) and equatorward wind (strong and variable south of and weak but persistent north of 30°S)¹⁷. The biogeographic break at 42°S is caused by the collision of the West Wind Drift Currents (Antarctic Circumpolar Current) with the Chilean coast, producing the northward Humboldt (Peru) Current and southward Cape Horn Current^{18,19}. These biogeographic breaks act as barriers for species with low dispersal^{20,21} and promote

¹Departamento de Ciencias Ecológicas, Facultad de Ciencias, Universidad de Chile, Santiago, Chile. ²Centro de Ecología y Manejo Sustentable (ESMOI), Coquimbo, Chile. ³Instituto de Ciencias Marinas y Limnológicas, (ICML), Laboratorio Costero de Recursos Acuáticos de Calfuco (LCRAC), Facultad de Ciencias, Universidad Austral de Chile, Valdivia, Chile. ⁴Centro FONDAF de Investigación de Dinámica de Ecosistemas Marinos de Altas Latitudes (IDEAL), Valdivia, Chile. ⁵Québec Océan, Département de Biologie, Université Laval, Québec, Canada. ✉email: dveliz@uchile.cl

| Sites | Coordinates | N Initial | N after filtering | AR | Hobs | Hexp | Hn.b | F _{IS} |
|------------|----------------------|-----------|-------------------|------|-------|-------|-------|-----------------|
| 2013–2014 | | | | | | | | |
| Tome | 36° 21' S; 72° 50' W | 23 | 18 | 1.5 | 0.120 | 0.134 | 0.138 | 0.192 |
| Valdivia | 39° 51' S; 73° 23' W | 23 | 17 | 1.48 | 0.099 | 0.131 | 0.135 | 0.279 |
| Ancúd | 41° 50' S; 73° 51' W | 22 | 9 | 1.41 | 0.084 | 0.124 | 0.132 | 0.380 |
| Quellón | 43° 08' S; 73° 36' W | 22 | 10 | 1.44 | 0.100 | 0.126 | 0.133 | 0.256 |
| 2020–2021 | | | | | | | | |
| Valparaíso | 33° 01' S; 71° 39' W | 28 | 28 | 1.56 | 0.131 | 0.137 | 0.139 | 0.061 |
| Dichato | 36° 31' S; 72° 57' W | 24 | 24 | 1.56 | 0.132 | 0.137 | 0.140 | 0.063 |
| Tumbes | 36° 38' S; 73° 05' W | 10 | 10 | 1.5 | 0.132 | 0.132 | 0.139 | 0.053 |
| Valdivia | 39° 51' S; 73° 23' W | 31 | 31 | 1.56 | 0.128 | 0.136 | 0.138 | 0.072 |
| Ancud | 41° 50' S; 73° 51' W | 31 | 31 | 1.54 | 0.122 | 0.134 | 0.136 | 0.108 |
| Quellón | 43° 08' S; 73° 36' W | 30 | 27 | 1.55 | 0.129 | 0.137 | 0.140 | 0.082 |
| Aysén | 45° 26' S; 72° 55' W | 30 | 29 | 1.54 | 0.124 | 0.135 | 0.137 | 0.093 |

Table 1. Summary of SNPs data of the crab *M. edwardsii* including sampling time (2013–2014 and 2020–2021), geographical coordinates, sample size before and after filtering, allelic richness (AR) observed heterozygosity (Hobs.), expected heterozygosity (Hexp.), expected heterozygosity corrected for sampling bias (Hn.b.) and F_{IS} at each study site.

non-random spatial movements of larvae in plankton, such as unidirectional movement²² or asymmetric dispersal among populations²³. Overall, these factors can impact the survival of the planktonic larvae²⁴, stochasticity in the recruitment of benthic species²⁵, and the substantial changes in allelic frequencies observed when different cohorts are analyzed²⁶.

Among the species whose distribution spans the 42°S biogeographic break (see Lancellotti and Vásquez (2000)¹⁶ for examples) is the mola rock crab *Metacarcinus edwardsii*, an important commercial fishery species that accounts for 75% of the Chilean artisanal cancrid fishery industry²⁷. *M. edwardsii* typically inhabits sandy, muddy, and gravel substrates up to 70 m deep in intertidal and subtidal zones. Its distribution stretches from Guayaquil (Ecuador) to the Strait of Magellan (Chile)²⁸, crossing the biogeographic breaks at 30° and 42°S.

Metacarcinus edwardsii attains sexual maturity at a carapace width of 100 mm for both sexes²⁹. Females mate with more than one male during the breeding season, but only one male becomes the genitor of the entire egg clutch^{30,31}. Hundreds of full-sib zoea are released from each female³². Megalopae are abundant during October and December, coinciding with the advection of warm waters towards the coast³³. Larvae recruit in association with the estuarine environment³⁴. Laboratory experiments have demonstrated a planktonic larval duration of 60 days at 15 °C³⁵, and a population genetic study based on eight microsatellites found high gene flow among populations over 700 km apart³⁶. However, the possible effect of the Humboldt Current on gene flow, temporal genetic stability, and population differentiation of the species across the biogeographic break at 42°S remains unknown.

In this study, we examine the spatial and temporal genetics of the *M. edwardsii* population using 4209 single nucleotide polymorphisms (SNPs). We explore the genetic differentiation, temporal stability, and gene flow patterns of *M. edwardsii* populations throughout their geographical distribution along the path of the Humboldt Current. Our study covers 1700 km of coastline, and we analyzed samples obtained in two different periods (2013–2014 and 2020–2021) on both sides of the biogeographic break at 42°S. Furthermore, we evaluate different migration models to infer the most probable gene flow patterns among *M. edwardsii* populations.

Results

SNP calling. From the 274 individuals collected, 20,230 raw SNPs were obtained. Following the removal of low-quality and outlier SNPs, 4209 SNPs and 234 individuals were retained for analysis. Many of the samples that failed genotyping were collected in 2013–2014, and this was likely due to poor DNA quality. The sample sizes before and after filtering are presented in Table 1.

Genetic diversity and population genetic analyses. Allelic richness (AR) and expected heterozygosity (Hexp) were similar across sites and time periods. Observed heterozygosity (Hobs) varied between 0.084 in Ancúd (2013) and 0.132 in Dichato and Tumbes (2021). Higher values of F_{IS} (inbreeding coefficient) were observed at sites with the smallest sample sizes (Table 1).

The population genetic analysis of all sites did not detect significant genetic differences among them. A principal coordinate analysis (PCoA) found individuals from all sites to overlap in the multivariate space, indicating no genetic differentiation among populations (Fig. 1). Pairwise F_{ST} values among sites were not statistically significant, except for comparisons between sites from the 2013–2014 period with small sample sizes and the other sites (Table 2). The Bayesian method implemented in STRUCTURE software found K = 1 to have the highest ln(K) value (mean LNP[K] = -476,972.5, P = 0.999). Barplots of higher values of K showed no evidence of differences among sites, even from different years. Therefore, K = 1 represented the biological model with the greatest support in our dataset (Fig. 2).

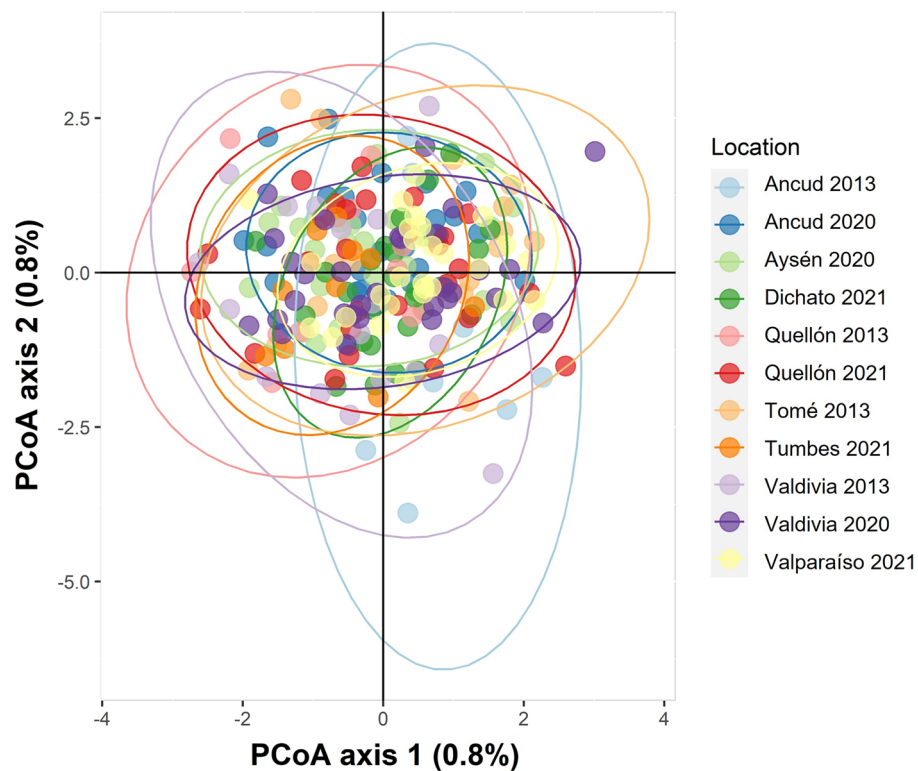


Figure 1. Principal coordinate analysis (PCoA) performed with *M. edwardsii*. The first and second Principal Components (x-axis and y-axis, respectively) capture 0.8% of the total variance each.

| | Valparaíso 2021 | Dichato 2021 | Tome 2014 | Tumbes 2021 | Valdivia 2013 | Valdivia 2020 | Ancud 2013 | Ancud 2020 | Quellón 2013 | Quellón 2021 | Aysén 2020 |
|-----------------|-----------------|--------------|-----------|-------------|---------------|---------------|------------|------------|--------------|--------------|------------|
| Valparaíso 2021 | | 0.002 | 0.002 | 0.002 | 0.003 | 0.001 | 0.009 | 0.001 | 0.005 | 0.001 | 0.001 |
| Dichato 2021 | 0.037 | | 0.002 | 0.001 | 0.002 | 0.001 | 0.007 | 0.002 | 0.002 | 0.002 | 0.001 |
| Tome 2014 | 0.051 | 0.057 | | 0.000 | 0.000 | 0.002 | 0.005 | 0.001 | 0.001 | 0.002 | 0.002 |
| Tumbes 2021 | 0.124 | 0.299 | 0.488 | | -0.003 | 0.001 | 0.006 | 0.001 | 0.000 | 0.001 | 0.002 |
| Valdivia 2013 | 0.018 | 0.051 | 0.493 | 0.976 | | 0.001 | 0.002 | 0.001 | 0.001 | 0.003 | 0.003 |
| Valdivia 2020 | 0.032 | 0.203 | 0.078 | 0.303 | 0.153 | | 0.009 | 0.000 | 0.004 | 0.001 | 0.001 |
| Ancud 2013 | 0.000 | 0.008 | 0.037 | 0.037 | 0.203 | 0.000 | | 0.007 | 0.002 | 0.008 | 0.009 |
| Ancud 2020 | 0.037 | 0.024 | 0.209 | 0.393 | 0.338 | 0.488 | 0.000 | | 0.003 | 0.001 | 0.001 |
| Quellón 2013 | 0.008 | 0.112 | 0.320 | 0.584 | 0.406 | 0.015 | 0.256 | 0.026 | | 0.004 | 0.004 |
| Quellón 2021 | 0.260 | 0.038 | 0.048 | 0.203 | 0.011 | 0.075 | 0.008 | 0.258 | 0.032 | | 0.000 |
| Aysén 2020 | 0.051 | 0.053 | 0.021 | 0.064 | 0.011 | 0.086 | 0.000 | 0.100 | 0.011 | 0.361 | |

Table 2. Pairwise FST (above diagonal) and corrected P-value (below diagonal) values for the sample sites.

Patterns of reciprocal migration among sampling sites. Among the models evaluated with MIGRATE software, the full model had the highest Bezier approximation scores for both the 2013–2014 and 2020–2021 datasets (Fig. 3). This result indicates that gene flow occurs between all sites. Migration rate estimates obtained with EEMS had $\log(m) \sim 0$ for all sites in the 2013–2014 and 2020–2021 datasets independently, indicating gene flow among sites (Fig. 4) and suggesting the absence of barriers to migration among sites. Finally, while a divMigrate analysis indicated strong gene flow among sites for both 2013–2014 and 2020–2021 (Fig. 5), bootstrap did not detect significant asymmetric gene flow among them ($P > 0.05$). Overall, the proportion of migrants between pairs of sites appeared to vary based on the type of analysis performed, all three methods used (MIGRATE, EEMS, and divMigrate) consistently indicated gene flow among all sites for both time periods.

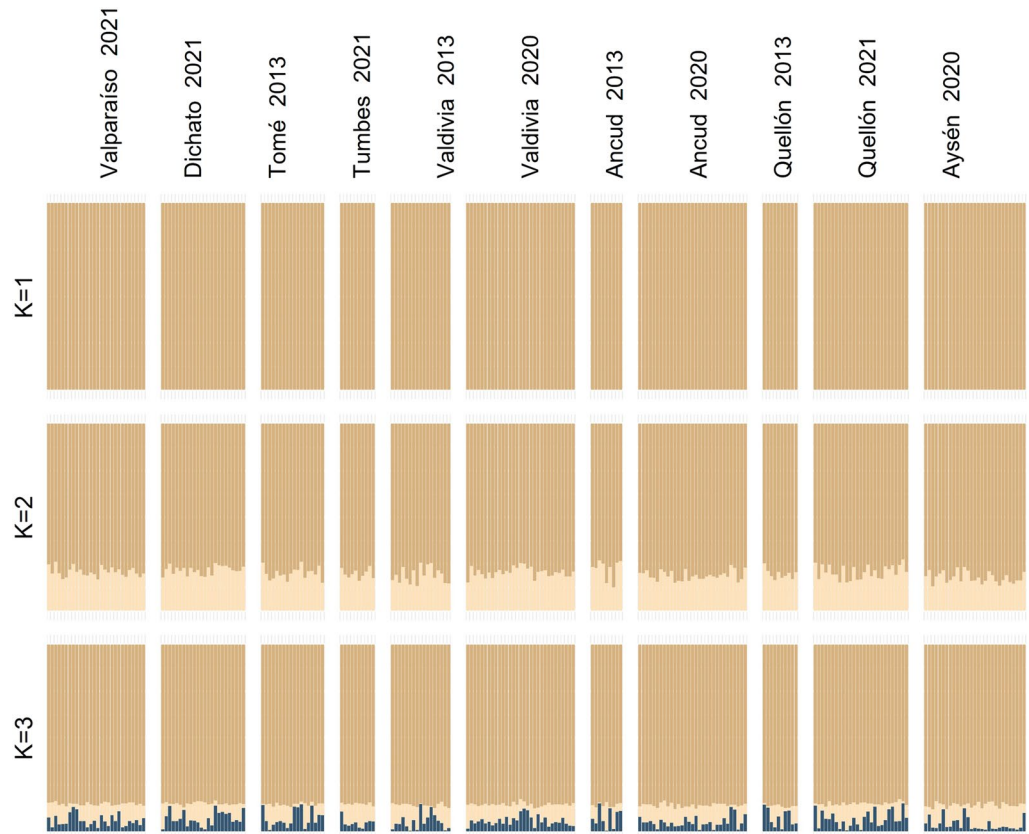


Figure 2. Population structure of the crab *M. edwardsii* inferred using the software STRUCTURE for $K=1$ to $K=3$ of the 234 individuals analyzed after filtering SNPs. A vertical bar represents each individual, and each color represents the probability of belonging to one of the K genetic clusters.

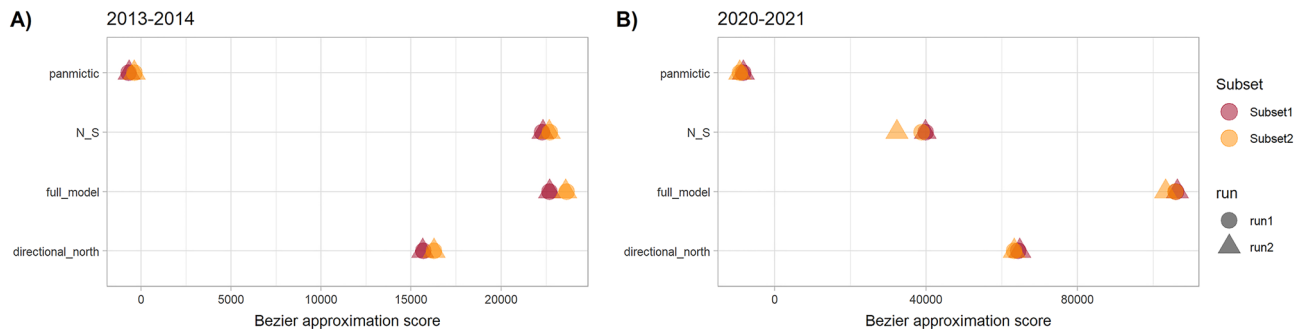


Figure 3. Values of the Bezier approximation score for the different models tested with MIGRATE software separately for 2013–2014 and 2020–2021. Panmictic = panmixia model; N_S = northbound migration at sites north of the biogeographic break and southbound at sites south of the biogeographic break; full model = full migration; directional_north = directional migration towards the north. The full model had the highest Bezier approximation value ($P > 0.999$) with both datasets.

Discussion

The results of this study show a clear pattern of high gene flow among *M. edwardsii* populations and temporal genetic stability across its geographical distribution. In addition, they indicate that the biogeographic break at 42°S and the Humboldt Current do not have a significant impact on gene flow among its populations. Compared to a previous study using eight microsatellites that sampled *M. edwardsii* from 700 km of the Chilean coastline, we found similar results in a broader geographical area.

Gene flow over large geographic areas has also been reported for other crustacean species with a similar duration of planktonic larval development (two to three months), for example, the Chilean mole crab *M. edwardsii* has a planktonic larval phase of 60 days³⁵, the brown crab *Cancer pagurus* of three months⁴¹, the Norway lobster *Nephrops norvegicus* of four to eight weeks⁴² and the blue crab *Callinectes sapidus* of five to ten

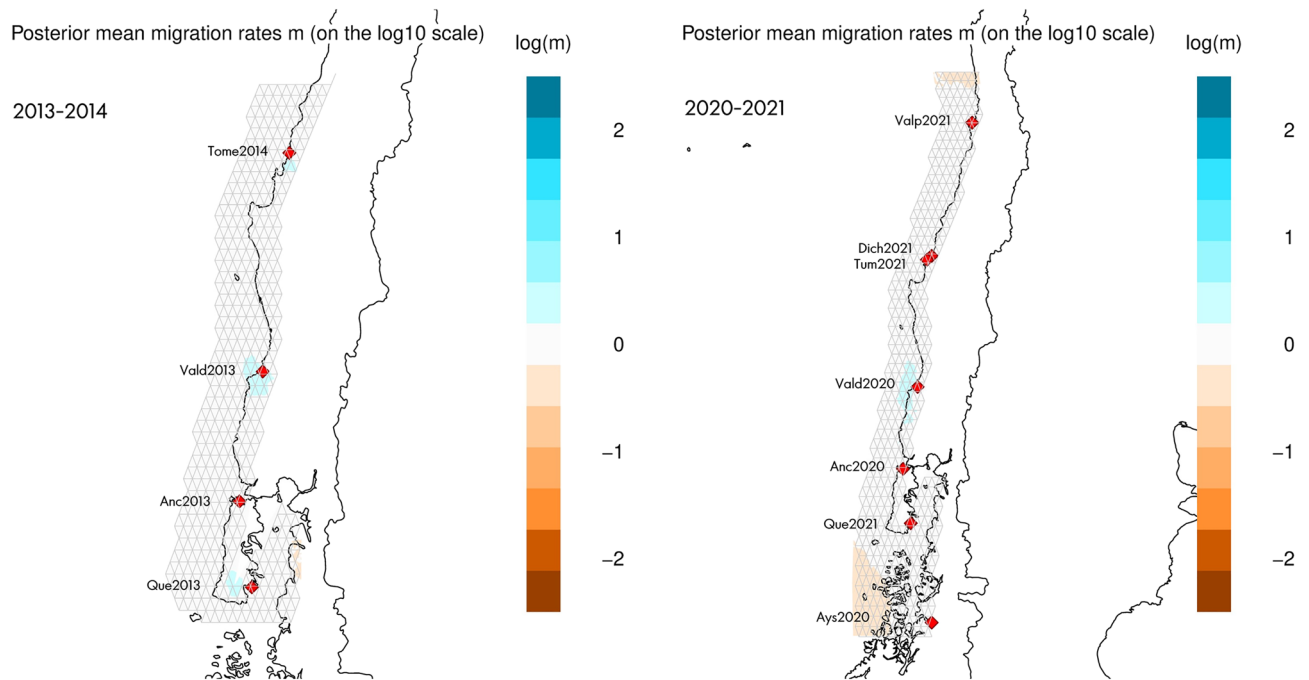


Figure 4. Effective Migration Rates for 2013–2014 (left) and 2020–2021 (right) estimated with the EMMS software³⁷. $\log(m)$ denotes the effective migration rate on a \log_{10} scale relative to the overall migration rate throughout the habitat. The blue colors represent areas where the effective migration is higher than average, while brown colors represent areas where effective migration is lower than average. Maps drawn using library rEEMSpots³⁸ implemented in R software³⁹.

weeks⁴³. For these species, population genetic analyses not showed statistical differences among populations of *M. edwardsii* separated by 1700 km (this study), populations of *C. pagurus* separated by 1300 km (from the Norwegian Sea to the Kattegat straits)⁴⁴, populations of *N. norvegicus* separated by 500 km in southern Iceland⁴² and populations of *C. sapidus* separated by 740 km in the Brazilian coast⁴⁵.

Biogeographic breaks generally produce discontinuity between populations located on either side. Evidence for this phenomenon has been observed in different coastal areas. For example, genetic differences were observed for gobiid fishes located on opposite sides of the Mona Passage in the Caribbean Sea⁴⁶, for different benthic invertebrates on the North and South islands of New Zealand⁴⁷ and the Southeast Australian Biogeographic Barrier⁴⁸, and for stomatopods located in Northern and Southern Indonesia⁴⁹. However, the biogeographic discontinuities are not impermeable for all species. Previous studies have demonstrated that two-thirds of benthic species with pelagic development do not present a genetic discontinuity between Alaska and California²¹ and show that populations of echinoderms and some crustaceans do not show discontinuity either side of the 30°S biogeographic break in Chile²⁰.

In the Chilean biogeographic break at 42°S, there is evidence of species with different propagule (zoospores or larvae) duration inhabiting both sides of the zone. The kelp species *Macrocystis pyrifera*, whose zoospores quickly settle a few meters from their parents⁵⁰, and the mytilid *Mytilus chilensis*, whose planktonic larval phase lasts between 20 and 45 days^{51,52}, show genetic differences between the two zones^{50,53}. However, for the commercial gastropod *Concholepas concholepas* that have a planktonic larval duration of three months⁵⁴, Cardenas et al.⁵⁵ did not detect genetic differences in populations located at north and south of the 42°S biogeographic break. Therefore, as would be expected for a species with the longest period of planktonic larval development, this break was found not to affect gene flow among its populations. Overall, the analysis suggest that the duration of the planktonic larval stage is an important aspect in the gene flow of species inhabiting both sides of a biogeographic break.

An interesting issue observed in *M. edwardsii* is the temporal stability of the genetic variability across time periods (2013–2014 and 2020–2021). The analyses did not detect population differentiation, except for slight but significant differences in F_{ST} when pairs of sites with small sample sizes were compared. This temporal stability has been previously described for this species; Rojas-Hernández et al.³⁶ did not detect differences in microsatellite variability in four cohorts of megalopas (over the period 2011 to 2014) collected at Los Molinos (39° 51' S; 73° 23' W). To our knowledge, there is little evidence of the temporal genetic stability of populations of benthic marine organisms. For example, temporal genetic stability was described in the crab *Carcinus maeneas* in the Iberian Peninsula^{56,57} and in the crab *Cancer pagurus* in Sweden⁴². Temporal genetic similarity of different adult cohorts was also studied in the Pacific Geoduck Clam *Panopea generous*⁵⁸ and the Arctic surf clam *Mactromeris polynyma*⁵⁹. Until now, there have been limited studies on temporal genetic stability as compared to those describing changes in allele frequencies between generations in other species of benthic marine organisms^{60,61}.

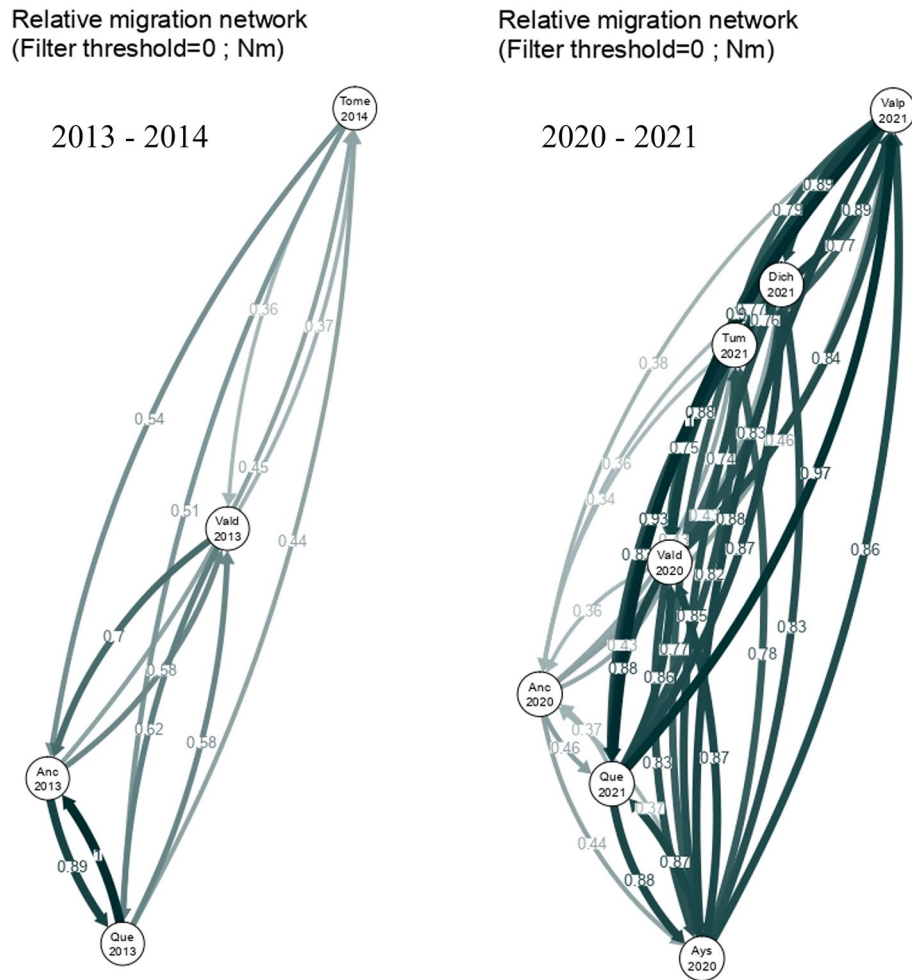


Figure 5. Migration network for 2013–2014 (left) and 2020–2021 (right) determined with divMigrate⁴⁰. Circles represent sampling sites, and each arrow represents the direction and magnitude (arrow edge value) of the relative migration between sites.

It is important to note that we expected a northward asymmetric gene flow following the Humboldt Current. However, the migration analyses suggest that gene flow is not unidirectional in this area. The different models tested with the Migrate software showed gene flow among all sites but with different proportions between pairs of sites. This observation is supported by EEMS and divMigrate analyses, which also indicated a variable number of migrants between pairs of sites. In other studies, clear patterns of asymmetry produced by global circulation have previously been described for different species. In the Southern Ocean, the ocean currents produce asymmetry in population connectivity of the shrimp *Nematocarcinus lanceopes*⁶². In addition, the Benguela Current drives asymmetry in gene flow in the cosmopolitan bluefish *Pomatomus saltatrix*⁶³, the Asia Northwestern Pacific Current in the brown seaweed *Sargassum fusiforme*⁶⁴, and the North Pacific Current in the sea cucumber *Parastichopus californicus*⁶⁵. Overall, our data suggest that the planktonic larval duration of *M. edwardsii* prevents the 42°S biogeographic break from negatively impacting in gene flow, consistent with the absence of spatial genetic differentiation in our results.

Finally, our results suggest that the duration of the planktonic larval stage is an important factor in determining the genetic structure of the species, allowing gene flow between distant populations despite the presence of a biogeographic break and currents that are primarily unidirectional. The findings of this study will be important for fishery management of *M. edwardsii* since artisanal fishing is partly dependent on this species in the south-central zone of Chile.

Materials and methods

Sampling sites. A total of 274 adult *M. edwardsii* crabs were collected for this study. Nations (1975)²⁸ described the geographical distribution of *M. edwardsii* from Guayaquil (Ecuador) to the Strait of Magellan (Chile). We searched for this species at different sites in its described geographic distribution in Chile however it was not found north of Valparaiso (32°S). Our sampling effort north of Valparaiso consisted of field work in Arica (18° 48' S), Pisagua (19° 35' S), Iquique (20° 12' S), Antofagasta (23° 65' S), Bahía Inglesa (27° 07' S), and Coquimbo (29°59'S) in Chile. Our observations highlight the need for current, up-to-date biogeographical

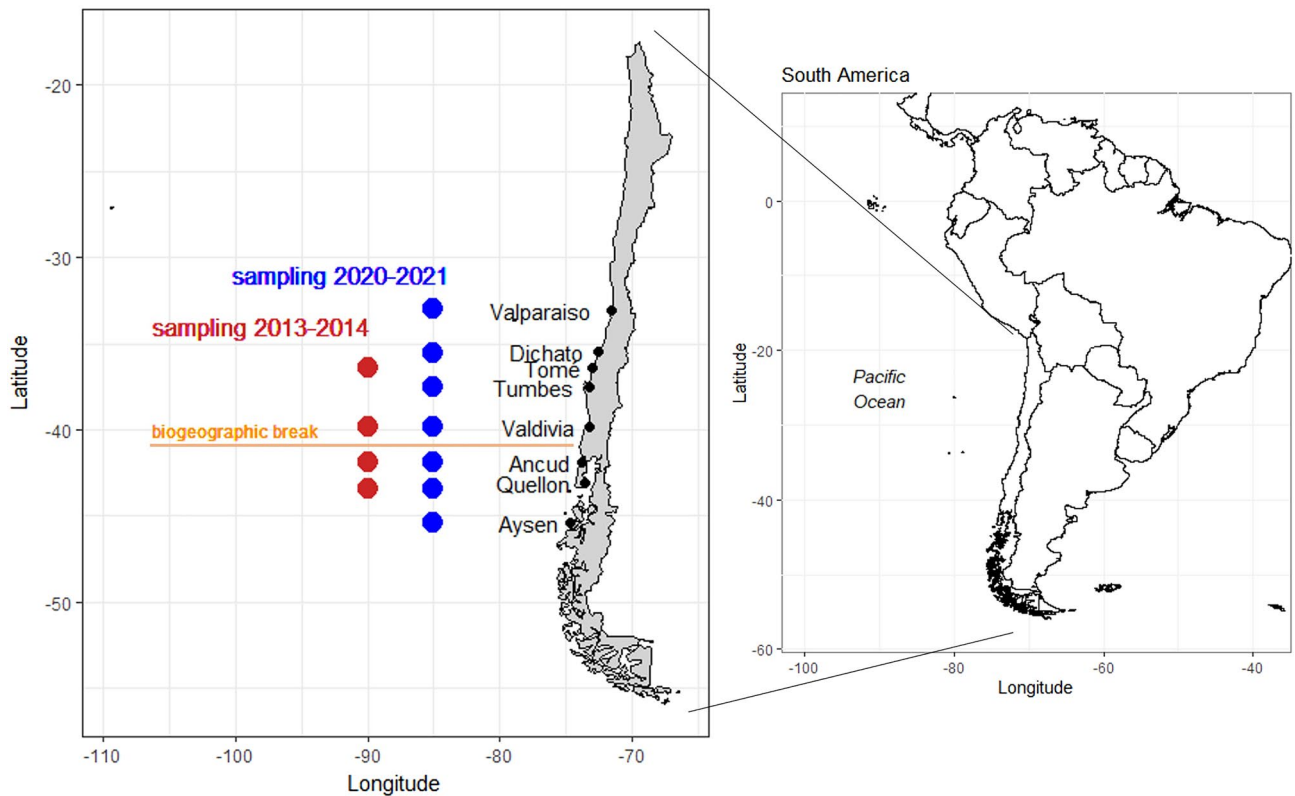


Figure 6. Sampling sites of *M. edwardsii* in Chile. Red points represent samples obtained in 2013–2014, while blue points represent samples obtained in 2020–2021. The horizontal line in orange represents the biogeographic break at 42°S. Maps drawn using library maps⁶⁶ and mapdata⁶⁷ implemented in R software⁶⁸.

distribution for this and other crab species. We used two groups of samples, first group collected in 2013–2014 and a second group collected in 2020–2021. The samples from 2013–2014 were reported by Rojas-Hernández et al.³⁶ that performed a population genetic analysis using eight microsatellites and covering sites along 700 km of the Chilean coast. The samples used here and reported in Rojas-Hernández et al.³⁶ were from Tomé (n = 23), Valdivia (n = 23), Ancud (n = 22), and Quellón (n = 22). The samples collected in 2020–2021 were collected in Valparaíso (n = 28), Dichato (n = 24), Tumbes (n = 10), Valdivia (n = 31), Ancud (n = 31), Quellón (n = 30), and Aysén (n = 30) (Fig. 6). Crabs were collected by local fishermen using commercial crab traps, and a pereopod from each specimen was stored in 95% ethanol until analysis.

Sequencing and SNP calling. For each crab, DNA was extracted from a small piece of muscle and subjected to massively parallel sequencing at Dart Diversity Arrays Technology Pty Ltd. (DART; Canberra, Australia). Following the methods of Kilian et al.³⁹, each DNA sample was digested using the restriction enzymes *PstI* and *HpaII*, and fragments > 200 bp were ligated with an 8 bp barcode prior to polymerase chain reaction (PCR) amplification. The PCR products were then standardized and sequenced using the Illumina HiSeq 2500 platform (San Diego, USA).

Sequences generated from each lane were processed using proprietary DART PL analytical pipelines. Barcode region and complete reads were filtered by quality parameters (barcode region: Phred > 30; complete read: Phred > 10). Using DART PL proprietary algorithm, identical sequences were collapsed into “fastcoll files” and low-quality bases from singleton tags were corrected into correct bases using collapsed tags with multiple members as a template. The proprietary DART PL algorithm (named DARTsoft14) was used to call the SNPs. For this study, call quality was assured by a high average read depth per locus (average of all markers was more than 10 reads/locus) and a sequencing volume per sample of 1.64 million reads. A total of 83 technical replicates of the 275 unique samples were run to estimate the reproducibility of the reported markers (> 99%). All this information was provided by DART PL. The resulting SNP data was filtered using the dartR library⁶⁹ in the R statistical software⁶⁸, retaining: (a) only one SNP in reads containing two or more SNPs, (b) SNPs with a read depth > 5 or < 100, and (c) SNPs with a > 95% reproducibility score, an index provided by DART PL that indicates the proportion of replicate technical samples with a consistent marker score (reproducible result). SNPs that were monomorphic, had > 15% missing data, or a minimum allele frequency (MAF) < 1% were removed, as were samples with > 15% missing data. Based upon a relatedness value $r > 0.4$, we detected a pair of related individuals collected in Concepción in 2014. One individual was removed to avoid potential bias in the population genetic estimation.

All SNPs showing signs of selection were removed to avoid potential bias in estimating the differences among sampling sites. Three different approaches were used here: (a) the likelihood-based method implemented by the *outflank* function of the *dartR* library in R, (b) the Bayesian method implemented in the *BayeScan* software⁷⁰, and (c) the relationship between F_{ST} and heterozygosity implemented in the *Fsthet* library⁷¹ in R software. We removed all SNPs that showed evidence of selection in any of these three approaches. Finally, SNPs with significant departures from Hardy–Weinberg equilibrium for one or more sites were removed using the *dartR* library in R, and SNPs with linkage disequilibrium > 0.2 in all sampling sites were removed using the *PLINK* 2.0 software⁴⁰.

Genetic diversity and population genetic structure. Genetic diversity at each sampling site was described with expected heterozygosity (H_{exp}), expected heterozygosity corrected for sampling bias ($H_{n.b.}$), observed heterozygosity (H_{obs}), and the inbreeding coefficient (F_{IS}) estimated using the *GENETIX* v 4.05 software⁷². Allelic richness (AR) was estimated using the *divBasic* function of the *diversity* library in R⁷².

Population genetic structure was estimated using the SNP database after removing all SNPs with signals of selection. Three population genetic methods were used with all data obtained in 2013–2014 and 2020–2021: (a) PCoA to describe the distribution of individuals in multivariate space, using the *dartR* library in R, (b) pairwise F_{ST} calculated using *GENETIX*⁷³ with 5000 permutations followed by the Bonferroni correction to the P-value, and (c) estimation of the most probable number of genetic clusters (K) using the Bayesian approach implemented in the *STRUCTURE* software⁷⁴. The admixture model and correlation of the allele frequencies were used as input. The procedure was performed three times for each K between 1 and 6 with a burn-in of 100,000 iterations and an after-burn-in of 200,000 iterations. The probability of each value of K was estimated as described in the *STRUCTURE* manual⁷⁵.

Patterns of reciprocal migration among sampling sites. Separately analyzing samples from periods 2013–2014 and 2020–2021, we estimated the direction and magnitude of the gene flow among sampling sites using three methods.

First, the historical migration rates was assessed using the *Migrate* software⁷⁶ and two random subsets of 1000 SNPs per period. This approach employs a coalescent method to estimate mutation-scaled migration rates (M) for each group over the last $4N_E$ generations. The Bayesian inference of the *Migrate* software was used with the default settings except for the following run options: (a) one single long run utilizing heating with temperatures of 1.0, 1.5, 3.0, and 1,000,000; (b) 1,000,000 genealogies were run with a sample increment of 10; and (c) the first 100,000 genealogies. The uniform prior distribution was used for Θ (from 0 to 0.1) and M (from 0 to 100,000). We tested four models that represent the probable gene flow in the studied area: (a) panmixia model, (b) full migration, (c) directional migration towards the north (considering northern variation in ocean circulation³⁷), and (d) northbound migration at sites north of the biogeographic break and southbound at sites south of the biogeographic break. To identify the best model, each analysis was performed twice for two independent subsets (a total of four runs per model), and we used the *bf.py* Python script provided by Beerli et al.³⁸ that compares the Bezier log marginal likelihood values obtained in each model. *MIGRATE* was run at the Cyber Infrastructure for Phylogenetic Research (CIPRES, www.phylo.org).

Second, the Estimating Effective Migration Surfaces (*EEMS*) software was used to visualize the gene flow patterns among sampling sites. *EEMS* estimates migration rates so that the genetic differences observed in the data match the genetic differences expected under an idealized stepping-stone model. These estimates are then interpolated across sampling sites to produce an “estimated effective migration surface”—a visual representation of genetic variation—that highlights regions with higher-than-average and lower-than-average historical gene flow⁷⁷. To capture the continuous population structure, *EEMS* covers the habitat with a dense regular grid, in which each deme exchanges migrants with its neighbors. If habitable regions are unsampled, estimates are based on the prior, which assumes no heterogeneity in migration rates⁷⁷. *EEMS* was run with the full set of 4209 SNPs using 500 demes and three independent chains of 5,000,000 MCMC iterations with a burn-in of 1,000,000 and sampling performed every 9999 iterations. The proposed variances were adjusted considering an acceptance rate ranging from 10 to 40%. Results were plotted using the *rEEMSpots* package⁷⁷ in R. Note that the results for effective migration rates are on a \log_{10} scale (denoted as $\log(m)$ in the plot) relative to the overall migration rate in the habitat. Thus, a $\log(m) = 1$ represents an effective migration ten times greater than the average, and a $\log(m) = -1$ corresponds to an effective migration ten times less than the average.

Finally, the direction and magnitude of the gene flow between pairs of sampling sites of *M. edwardsii* were estimated using the *divMigrate* function of the *diversity* library⁷² in R. The *Alcala's* statistic (Nm_{Alcala}) was used as a distance measurement, which incorporates information from both G_{st} and D and maybe generally better suited to different demographic scenarios⁷². Furthermore, the asymmetry of gene flow between pairs of sites was tested with the full set of 4209 SNPs using a bootstrap of 1000 iterations with Nm_{Alcala} used as a distance measure.

Data availability

The datasets generated and analysed during the current study are available in the *datos.uchile.cl* repository, <https://doi.org/10.34691/FK2/VF1R0K>. Further, raw data is available in GenBank with the following NCBI data accession: BioProject ID PRJNA863944 and BioSample accessions: SAMN30070490—SAMN30070763.

Received: 7 April 2022; Accepted: 10 August 2022

Published online: 20 August 2022

References

1. Thorson, G. Reproductive and larval ecology of marine bottom invertebrates. *Biol. Rev.* **25**, 1–45 (1950).
2. Weersing, K. & Toonen, R. J. Population genetics, larval dispersal, and connectivity in marine systems. *Mar. Ecol. Progr. Ser.* **393**, 1–12 (2009).
3. Hedgecock, D. Is gene flow from pelagic larval dispersal important in the adaptation and evolution of marine invertebrates?. *Bull. Mar. Sci.* **39**, 550–564 (1986).
4. Jenkins, S. R. & Hawkins, S. J. Barnacle larval supply to sheltered rocky shores: a limiting factor?. *Hydrobiologia* **503**, 143–151 (2003).
5. Pineda, J., Hare, J. A. & Sponaugle, S. Consequences for population connectivity. *Oceanography* **20**, 22–39 (2007).
6. Shanks, A. L. Mechanisms of cross-shelf dispersal of larval invertebrates and fish. In *Ecology of Marine Invertebrate Larvae* (ed. McEdward, L. R.) 324–367 (CRC, Boca Raton, 1995).
7. Shanks, A. L. Pelagic larval duration and dispersal distance revisited. *Biol. Bull.* **216**, 373–385 (2009).
8. Bradford, R. W., Griffin, D. & Bruce, B. D. Estimating the duration of the pelagic phyllosoma phase of the southern rock lobster, *Jasus edwardsii* (Hutton). *Mar. Freshw. Res.* **66**, 213–219 (2015).
9. Mileikovsky, S. A. Speed of active movement of pelagic larvae of marine bottom invertebrates and their ability to regulate their vertical position. *Mar. Biol.* **23**, 11–17 (1973).
10. Garrison, L. P. Vertical migration behavior and larval transport in brachyuran crabs. *Mar. Ecol. Progr. Ser.* **176**, 103–113 (1999).
11. Morgan, S. G. & Fisher, J. L. Larval behavior regulates nearshore retention and offshore migration in an upwelling shadow and along the open coast. *Mar. Ecol. Progr. Ser.* **404**, 109–126 (2010).
12. Cowen, R. K. & Castro, L. R. Relation of coral reef fish larval distributions to island scale circulation around Barbados, west indies. *Bull. Mar. Sci.* **54**, 228–224 (1994).
13. Rudorff, C. A. G., Lorenzetti, J. A., Gherardia, D. F. M. & Lins-Oliveira, J. E. Modeling spiny lobster larval dispersion in the Tropical Atlantic. *Fish. Res.* **96**, 206–215 (2009).
14. Allee, W. C. Studies in marine ecology. IV. The effect of temperature in limiting the geographic range of invertebrates of the Woods Hole littoral. *Ecology* **4**, 341–354 (1923).
15. Burton, R. S. Intraspecific phylogeography across the Point Conception biogeographic boundary. *Evolution* **52**, 734–745 (1998).
16. Lancellotti, D. A. & Vasquez, J. A. Biogeographical patterns of benthic macroinvertebrates in the southeastern Pacific littoral. *J. Biogeogr.* **26**, 1001–1006 (1999).
17. Hormazabal, S., Shaffer, G. & Leth, O. Coastal transition zone off Chile. *J. Geophys. Res.* **109**, C01021 (2004).
18. Mcdonald, A. M. The global ocean circulation: a hydrographic estimate and regional analysis. *Prog. Oceanogr.* **41**, 281–382 (1998).
19. Montecino, V. & Lange, C. B. The Humboldt Current System: Ecosystem components and processes, fisheries, and sediment studies. *Prog. Oceanogr.* **83**, 65–79 (2009).
20. Haye, P. A. *et al.* Phylogeographic structure in benthic marine invertebrates of the southeast Pacific Coast of Chile with differing dispersal potential. *PLoS ONE* **9**, e88613 (2014).
21. Kelly, R. P. & Palumbi, S. R. Genetic structure among 50 species of the northeastern Pacific rocky intertidal community. *PLoS ONE* **5**, e8594 (2010).
22. Gaylord, B. & Gaines, S. D. Temperature or transport? Range limits in marine species mediated solely by flow. *Am. Nat.* **155**, 769–789 (2000).
23. Wares, J. P., Gaines, S. D. & Cunningham, C. W. A comparative study of asymmetric migration events across a marine biogeographic boundary. *Evolution* **55**, 295–306 (2001).
24. Rumrill, S. S. Natural mortality of marine invertebrate larvae. *Ophelia* **32**, 163–198 (1990).
25. Jenkins, S. R., Marshall, D. & Fraschetti, S. Settlement and recruitment. In *Marine Hard Bottom Communities, Ecological Studies* Vol. 206 (ed. Wahl, M.) 177–190 (Springer, Berlin, 2009).
26. Marino, I. A. M. *et al.* Genetic heterogeneity in populations of the Mediterranean shore crab, *Carcinus aestuarii* (Decapoda, Portunidae), from the Venice Lagoon. *Estuar. Coast. Shelf. Sci.* **87**, 135–144 (2010).
27. Sernapesca. Estadística de pesca de Chile. <http://www.sernapesca.cl/informes/estadisticas> (2022).
28. Nation JD (1975) The Genus *Cancer*: Crustacea: Brachyura): Systematics, biogeography and fossil record. *Nat. Hist. Mus. Los Angeles County Sci. Bull.* **23** (1975).
29. Pardo, L. M., Fuentes, J. P., Olguin, A. & Orensanz, J. M. L. Reproductive maturity in the edible Chilean crab *Cancer edwardsii*: methodological and management considerations. *J. Mar. Biol. Assoc. U. K.* **89**, 1627–1634 (2009).
30. Rojas-Hernández, N., Veliz, D. & Pardo, L. M. Use of novel microsatellite markers for population and paternity analysis in the commercially important crab *Metacarcinus edwardsii*. *Mar. Biol. Res.* **10**, 839–844 (2014).
31. Pardo, L. M., Riveros, M. P., Fuentes, J. P., Rojas-Hernández, N. & Veliz, D. An effective sperm competition avoidance strategy in crabs drives genetic monogamy despite evidence of polyandry. *Behav. Ecol. Sociobiol.* **70**, 73–81 (2016).
32. Pardo, L. M. *et al.* High fishing intensity reduces females' sperm reserve and brood fecundity in a eubrachyuran crab subject to sex- and size biased harvest. *ICES J. Mar. Sci.* **74**, 2459–2469 (2017).
33. Pardo, L. M., Mora-Vásquez, P. & Garcés-Vargas, J. Asentamiento diario de megalopas de jaibas del género *Cancer* en un estuario micromareal. *Lat. Am. J. Aquat. Res.* **40**, 142–152 (2012).
34. Pardo, L. M., Rubilar, P. R. & Fuentes, J. P. North Patagonian estuaries appear to function as nursery habitats for marble crab (*Metacarcinus edwardsii*). *Reg. Stud. Mar. Sci.* **36**, 101315 (2020).
35. Quintana, R. Larval development of the Edible crab, *Cancer edwardsii* Bell, 1835 under laboratory conditions (Decapoda, Brachyura). *Rep. USA Mar. Biol. Inst.* **5**, 1–19 (1983).
36. Rojas-Hernández, N., Veliz, D., Riveros, M. P., Fuentes, J. P. & Pardo, L. M. Highly connected populations and temporal stability in allelic frequencies of a harvested crab from southern Pacific. *PLoS ONE* **11**, e0166029 (2016).
37. Strub, P. T., James, C., Montecino, V., Rutllant, J. A. & Blanco, J. L. Ocean circulation along the southern Chile transition region (38°–46°S): Mean, seasonal and interannual variability, with a focus on 2014–2016. *Prog. Oceanogr.* **172**, 159–198 (2019).
38. Beerli, P., Mashayekhi, S., Sadeghi, M., Khodaei, M. & Shaw, K. Population genetic inference with MIGRATE. *Curr. Protoc. Bioinform.* **68**, e87 (2019).
39. Kilian, A. *et al.* Diversity arrays technology: A generic genome profiling technology on open platforms. *Methods Mol. Biol.* **888**, 67–89 (2012).
40. Chang, C. C. *et al.* Second-generation PLINK: Rising to the challenge of larger and richer datasets. *GigaScience* **4**, 7 (2015).
41. Weiss, M. *et al.* Influence of temperature on the larval development of the edible crab, *Cancer pagurus*. *J. Mar. Biol. Assoc. UK* **89**, 753–759 (2009).
42. Pampoulie, C. *et al.* A pilot genetic study reveals the absence of spatial genetic structure in Norway lobster (*Nephrops norvegicus*) on fishing grounds in Icelandic waters. *ICES J. Mar. Sci.* **68**, 20–25 (2011).
43. Costlow, J. D. J. & Bookhout, C. G. The larval development of *Callinectes sapidus* Rathbun reared in the laboratory. *Biol. Bull.* **116**, 373–396 (1959).
44. Ungfors, A., McKeown, N. J., Shaw, P. W. & Andre, C. Lack of spatial genetic variation in the edible crab (*Cancer pagurus*) in the Kattegat - Skagerrak area. *ICES J. Mar. Sci.* **66**, 462–469 (2009).

45. Lacerda, A. L. F. *et al.* High connectivity among blue crab (*Callinectes sapidus*) populations in the Western South Atlantic. *PLoS ONE* **11**, e0153124 (2016).
46. Taylor, M. S. & Hellberg, M. E. Comparative phylogeography in a genus of coral reef fishes: biogeographic and genetic concordance in the Caribbean. *Mol. Ecol.* **15**, 695–707 (2006).
47. Arranz, V., Fewster, R. M. & Lavery, S. D. Geographic concordance of genetic barriers in New Zealand coastal marine species. *Aquat. Conserv. Mar. Freshw. Ecosyst.* **31**, 3607–3625 (2021).
48. Ayre, D. J., Minchinton, T. E. & Perrin, C. Does life history predict past and current connectivity for rocky intertidal invertebrates across a marine biogeographic barrier?. *Mol. Ecol.* **18**, 1887–1903 (2009).
49. Barber, P. H., Erdmann, M. V. & Palumbi, S. R. Comparative phylogeography of three codistributed stomatopods: origins and timing of regional lineage diversification in the coral triangle. *Evolution* **60**, 1825–1839 (2006).
50. Macaya, E. C. & Zuccarello, G. C. Genetic structure of the giant kelp *Macrocystis pyrifera* along the southeastern Pacific. *Mar. Ecol. Progr. Ser.* **420**, 103–112 (2010).
51. Ruiz, M., Tarifeño, E., Llanos-Rivera, A., Padget, C. & Campos, B. Efecto de la temperatura en el desarrollo embrionario y larval del mejillón, *Mytilus galloprovincialis* (Lamarck 1819). *Rev. Biol. Mar. Oceanogr.* **43**, 51–61 (2008).
52. Toro, J. E., Castro, G. C., Ojeda, J. A. & Vergara, A. M. Allozymic variation and differentiation in the Chilean blue mussel, *Mytilus chilensis*, along its natural distribution. *Genet. Mol. Biol.* **29**, 174–179 (2006).
53. Araneda, C., Larrain, M. A., Hecht, B. & Narum, S. Adaptive genetic variation distinguishes Chilean blue mussels (*Mytilus chilensis*) from different marine environments. *Ecol. Evol.* **6**, 3632–3644 (2016).
54. Disalvo, L. H. Observations on the larval and post-metamorphic life of *Concholepas concholepas* (Bruguière, 1789) in laboratory culture. *Veliger* **30**, 358–368 (1988).
55. Cardenas, L., Castilla, J. C. & Viard, F. Hierarchical analysis of the population genetic structure in *Concholepas concholepas*, a marine mollusk with a long-lived dispersive larva. *Mar. Ecol.* **37**, 359–369 (2016).
56. Domingues, C. P., Creer, S., Taylor, M. I., Queiroga, H. & Carvalho, G. R. Genetic structure of *Carcinus maenas* within its native range: larval dispersal and oceanographic variability. *Mar. Ecol. Progr. Ser.* **410**, 111–123 (2010).
57. Domingues, C. P., Creer, S., Taylor, M. I., Queiroga, H. & Carvalho, G. R. Temporal genetic homogeneity among shore crab (*Carcinus maenas*) larval events supplied to an estuarine system on the Portuguese northwest coast. *Heredity* **106**, 832–840 (2011).
58. Vadopalas, B., Pietsch, T. & Friedman, C. The proper name for the geoduck: resurrection of *Panopea generosa* Gould, 1850, from the synonymy of *Panopea abrupta* (Conrad, 1849) (Bivalvia: Myoidia: Hiattellidae). *Malacologia* **52**, 169–173 (2010).
59. Cassista, M. C. & Hart, M. W. Spatial and temporal genetic homogeneity in the Arctic surfclam (*Mactromeris polynyma*). *Mar. Biol.* **152**, 569–579 (2007).
60. Li, G. & Hedgecock, D. Genetic heterogeneity, detected by PCR-SSCP, among samples of larval Pacific oysters (*Crassostrea gigas*) supports the hypothesis of large variance in reproductive success. *Can. J. Fish. Aquat. Sci.* **55**, 1025–1033 (1998).
61. Schmidt, P. S., Phifer-Rixey, M., Taylor, G. M. & Christner, J. Genetic heterogeneity among intertidal habitats in the flat periwinkle, *Littorina obtusata*. *Mol. Ecol.* **16**, 2393–2404 (2007).
62. Dambach, J., Raupach, M. J., Leese, F., Schwarzer, J. & Engler, J. O. Ocean currents determine functional connectivity in an Antarctic deep-sea shrimp. *Mar. Ecol.* **37**, 1336–1344 (2016).
63. Reid, K. *et al.* Secondary contact and asymmetrical gene flow in a cosmopolitan marine fish across the Benguela upwelling zone. *Heredity* **117**, 307–315 (2016).
64. Hu, Z.-M., Zhang, J., Lopez-Bautista, J. & Duan, D.-L. Asymmetric genetic exchange in the brown seaweed *Sargassum fusiforme* (Phaeophyceae) driven by oceanic currents. *Mar. Biol.* **160**, 1407–1414 (2013).
65. Xuereb, A. *et al.* Asymmetric oceanographic processes mediate connectivity and population genetic structure, as revealed by RADseq, in a highly dispersive marine invertebrate (*Parastichopus californicus*). *Mol. Ecol.* **27**, 2347–2364 (2018).
66. Becker, R. A. & Wilks, A. R. R version by Ray Brownrigg. mapdata: Extra Map Databases. R package version 2.3.0. (2018b).
67. Becker, R. A. & Wilks, A. R. R version by Ray Brownrigg. Enhancements by TP Minka and A Deckmyn.maps: Draw Geographical Maps. R package version 3.3.0. <https://CRAN.R-project.org/package=maps> (2018a).
68. R Core Team. *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, Vienna, 2022).
69. Grube, B., Unmack, P. J., Berry, O. F. & Georges, A. dartr: An R package to facilitate analysis of SNP data generated from reduced representation genome sequencing. *Mol. Ecol. Resour.* **18**, 691–699 (2018).
70. Foll, M. & Gaggiotti, O. E. A genome scan method to identify selected loci appropriate for both dominant and codominant markers: A Bayesian perspective. *Genetics* **180**, 977–993 (2008).
71. Flanagan, S. P. & Jones, A. G. Constraints on the Fst-heterozygosity outlier approach. *J. Hered.* **108**, 561–573 (2017).
72. Keenan, K., McGinnity, P., Cross, T. F., Crozier, W. W. & Prodohl, P. A. diveRsity: An R package for the estimation and exploration of population genetics parameters and their associated errors. *Methods Ecol Evol* **4**, 782–788 (2013).
73. Belkhir, K., Borsa, P., Chikhi, L., Raufaste, N. & Bonhomme, F. GENETIX 4.05, Logiciel sous Windows pour la Genetique des Populations. Laboratoire Genome, Populations, Interactions, CNRS UMR 5000 (Université de Montpellier II, Montpellier, France, 2000).
74. Pritchard, J. K., Stephens, M. & Donnelly, P. Inference of population structure using multilocus genotype data. *Genetics* **155**, 945–959 (2000).
75. Pritchard, J. K., Wen, X. & Falush, D. *Documentation for Structure Software: Version 2.3*. University of Oxford <http://pritch.bsd.uchicago.edu/structure.html> (2010).
76. Beerli, P. & Felsenstein, J. Maximum likelihood estimation of a migration matrix and effective population sizes in n subpopulations by using a coalescent approach. *Proc. Natl. Acad. Sci. U.S.A.* **98**, 4563–4568 (2001).
77. Petkova, D., Novembre, J. & Stephens, M. Visualizing spatial population structure with estimated effective migration surfaces. *Nat. Genet.* **48**, 94–100 (2016).

Acknowledgements

Thanks to JP Fuentes, M Riveros, A Auger, MJ Bruning, C Asmut, M. Gallardo for helping with sampling. This research was financed by ANID Fondecyt 1200589. LMP the research center FONDAP Dynamics of High Latitude Marine Ecosystems (IDEAL), grant N°15150003.

Author contributions

D.V., C.V.R. and L.M.P. conceived and designed the study. I.G., D.V. and L.M.P. sampled some sites. D.V., N.R.H. and C.Z. analyzed the data. D.V., N.R.H., C.V.R., C.Z., I.G., L.M.P. drafted the manuscript. All authors read, approved, and contributed to the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to D.V.

Reprints and permissions information is available at www.nature.com/reprints.

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



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

© The Author(s) 2022