

Research Article

Genetic population structure of the pleasure oyster *Crassostrea corteziensis* (Hertlein, 1951) on the east coast of the Gulf of California

Noé Díaz-Viloria¹, Ana María Ibarra², Claudia Gabriela Rivera-Apodaca¹
Nicole Reguera-Rouzaud¹, Norma Karina Hernández-Ibarra¹ & Pedro Cruz-Hernández²

¹Instituto Politécnico Nacional, Centro Interdisciplinario de Ciencias Marinas
La Paz, B.C.S., México

²Laboratorio de Genética Acuícola, Centro de Investigaciones Biológicas del Noroeste
La Paz, B.C.S., México

Corresponding author: Pedro Cruz-Hernández (pcruz@cibnor.mx)

ABSTRACT. The pleasure oyster, *Crassostrea corteziensis*, is an endemic species with high demand in northwestern México. Its incipient fishery with reduced captures makes it a good candidate for implementing management strategies to improve natural stocks and aquaculture practices. In this study, genetic diversity and population structure of *C. corteziensis* from four locations on the east coast of the Gulf of California (Bahía Lobos at Sonora, Las Glorias and Bahía de Ceuta at Sinaloa, and Boca de Camichín at Nayarit) were examined. Analysis with six microsatellites showed a high genetic diversity; however, null alleles were detected in almost all locations. After excluding null alleles and the locus *Ccor21*, a non-panmictic population of *C. corteziensis* was found along the coastline of the study area. Several factors could be contributing to restricted genetic connectivity among locations in the region without clear isolation by distance pattern, like random genetic drift maintained by larval retention, variance in reproductive success, asymmetric and restricted gene flow, and probably departures in the sex ratio from 1:1 and differences in the habitat and local adaptations to environmental conditions within each location.

Keywords: microsatellites; non-panmictic population; null alleles; genetic diversity; aquaculture practices, random genetic drift

INTRODUCTION

The pleasure oyster, *Crassostrea corteziensis* (Hertlein, 1951), occurs from the Gulf of California to Peru, shows high fecundity, a larval period of 18-22 days, and at least two strong spawning periods: one during the summer and the second in autumn (Stuardo & Martínez 1975, Chávez-Villalba et al. 2005, Rodríguez-Jaramillo et al. 2008).

The Gulf of California is a semi-enclosed sea and experiences seasonal reversal circulation. In summer, the surface circulation in the Gulf of California is characterized by a cyclonic pattern, with an eddy in the northern gulf. Additionally, a northwestward coastal

current along the mainland side extends the southward inflowing current. This coastal current plays a crucial role in transporting nutrients and larvae towards the northern region of the gulf (Lavín & Marinone 2003, Marinone 2003, 2012). During autumn and winter, the main circulation changes to anticyclonic, reversing the northern gulf eddy and the coastal current in the Gulf of California (Marinone 2003).

Oysters are the engineers of ecosystems like estuaries because they provide habitat for the entire ecosystem and remove suspended sediments (Silliman 2018, Lapègue et al. 2023). *C. corteziensis* has traditionally been harvested by local fishermen along the coasts of the Sonora, Sinaloa, and Nayarit Mexican

states. This native species has commercial value and aquaculture potential in México and Central America (Pérez-Enríquez et al. 2008, Enríquez-Espinoza & Grijalva-Chon 2010). This species has been cultured in Nayarit for over 25 years (Stuardo & Martínez 1975). Conversely, commercial captures of the pleasure oyster's wild populations have decreased, partly because of heavy exploitation in Sonora and Sinaloa and a reduction in freshwater flow into estuaries (Chávez-Villalba et al. 2005).

The first study in *C. corteziensis* reported differences in electrophoretic patterns of protein bands between individuals from two populations at Nayarit, México, under brackish (a mixture of fresh water and seawater) and hypersaline conditions. However, genetic population differentiation was not supported (Rodríguez-Romero et al. 1988). No genetic differences among *C. corteziensis* embankments from four locations (Guaymas, Sonora; Bahía de Ceuta, Sinaloa; north Nayarit, and south Nayarit) were found in México with four polymorphic allozyme loci (Pérez-Enríquez et al. 2008). However, allozymes reveal a limited proportion of the total variation and might be subjected to natural selection (Hellberg et al. 2002).

Several studies in different oyster species, using microsatellites and SNPs, have found genetic population structure due to the higher polymorphism than allozymes and the highest number of nuclear markers, respectively. In the flat oyster (*Ostrea edulis*), three population groups (Atlantic, western Mediterranean, and eastern Mediterranean) were identified with five microsatellites (Launey et al. 2002). In *Crassostrea virginica*, significant levels of geographic differentiation were found at Chesapeake Bay (Rose et al. 2006). Also, in *C. virginica*, a modest population structure with restricted gene flow among populations from six lagoons along the coast of Veracruz, México, was reported with five microsatellite loci (Galindo-Sánchez et al. 2008). A study with SNPs in *Ostrea lurida* reported genetic population differences with neutral and outlier loci; such differences between six regions from the west coast of North America agreed with biogeographic provinces to a good extent (Silliman 2018). Two recent studies in *Saccostrea echinata* and *O. edulis* with SNPs showed four genetic groups across northern Australia and New Caledonia, and through the North Sea, Atlantic Ocean, the western part of the Mediterranean Sea and the eastern part of the Mediterranean Sea and Black Sea, respectively (Nowland et al. 2019, Lapègue et al. 2023). Results in *O. edulis* followed those obtained with microsatellites by Launey et al. (2002).

Still, the study with SNPs found genetic differences in one population group from the Atlantic, being two different genetic groups (Lapègue et al. 2023).

All studies that showed the genetic population structure of wild populations from different oyster species suggested that possible factors influencing differentiation were local gene flow (Launey et al. 2002, Rose et al. 2006), short pelagic phase, isolation by distance (Launey et al. 2002, Rose et al. 2006), physical barriers, oceanographic barriers (oceanic fronts, temperature, salinity, and pathogen presence), asymmetric gene flow due to direction of currents, larval retention within a bay during reproductive season, bottlenecks, local adaptations, and genetic barriers (Galindo-Sánchez et al. 2008, Silliman 2018, Nowland et al. 2019, Lapègue et al. 2023).

At the moment, all previous genetic population studies in *C. corteziensis* were carried out with allozymes, and no differentiation was found between populations using these markers. Otherwise, microsatellite markers have been useful for detecting significant levels of genetic population differentiation in other oyster species. In agreement with this, the objective of this study was to assess the genetic diversity and population structure of *C. corteziensis* in four locations along the eastern coast of the Gulf of California using microsatellite markers for the first time in this species, analyzing a total of six microsatellites.

MATERIALS AND METHODS

Sample collection

From 40 to 50 wild specimens of *C. corteziensis* per location were collected in the northwestern region of México: Bahía de Lobos, Sonora (27°20'N, 110°30'W), Bahía de Ceuta, Sinaloa (24°05'N, 107°09'W), and Boca de Camichín, Nayarit (21°45'N, 105°30'W) during January 2004 and March 2005, Las Glorias, Sinaloa (25°17'N, 108°31'W) was collected in 2007 (Fig. 1). Three of these four samples (Bahía de Lobos, Bahía Ceuta, and Boca de Camichín) were collected by Pérez-Enríquez et al. (2008). Pérez-Enríquez and coworkers employed molecular identification based on partial 16srRNA gene sequences and confirmed that individuals of these three localities represent *C. corteziensis*. The oysters were transported alive or frozen to the Aquaculture Genetics Laboratory at the Centro de Investigaciones Biológicas del Noroeste (CIBNOR), where a small piece of adductor muscle from each individual was removed, preserved in ethanol (95%), and kept at 4°C until DNA extraction.

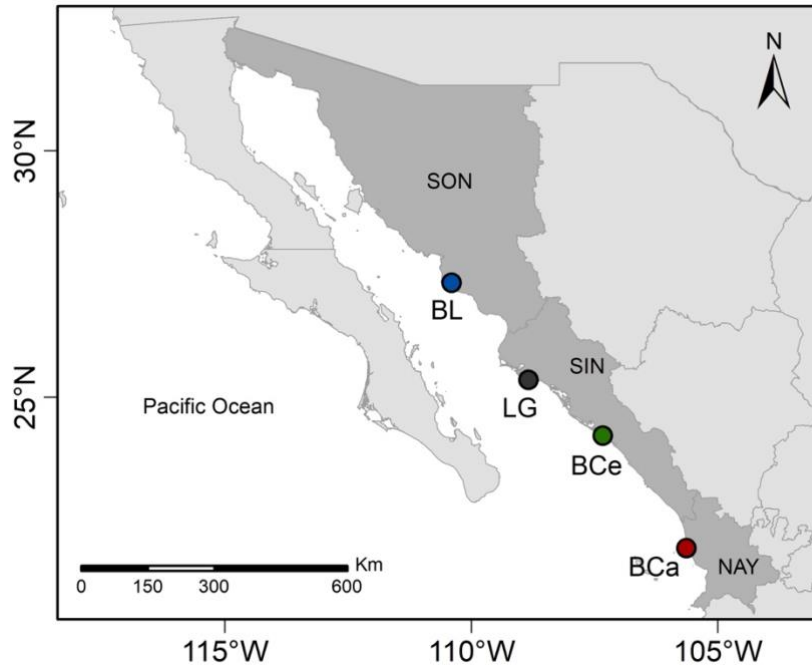


Figure 1. Collection sites of *Crassostrea corteziensis*. BL: Bahía de Lobos, SON: Sonora, LG: Las Glorias, BCe: Bahía de Ceuta, SIN: Sinaloa, BCa: Boca de Camichín, and NAY: Nayarit.

Microsatellite analysis

Genomic DNA from each individual was extracted in 2007 with Wizard® SV 96 Genomic DNA Purification System (Promega, Madison, WI). Gel electrophoresis evaluated DNA extractions, and the purity was assessed in a spectrophotometer ($A_{260}/A_{280} \sim 1.8$). A total of 186 individuals were examined with six polymorphic loci (*Ccor8*, *Ccor10*, *Ccor11*, *Ccor21*, *Ccor26*, and *Ccor28*) (Cruz et al. 2007) (Table 1). DNA amplifications were carried out in a volume of 6 μ L, containing 1 μ L of DNA, 0.18 μ M of $MgCl_2$, 0.12 mM of each dNTP, 0.025 units of Taq polymerase, and 1X Taq buffer (Invitrogen, Waltham, MA). The polymerase chain reaction (PCR) thermal conditions in iCycler (BIO-RAD, Hercules, CA) or 2720 Thermal Cycler (Applied Biosystems, Carlsbad, CA) were 5 min at 95°C, followed by 40 cycles of 30 s at 95°C, 30 s at 50°C, and 1 min at 72°C, and a final step of 15 min at 4°C (Cruz et al. 2007). PCR products were size screened on 7.5 M urea, 5% polyacrylamide gels and visualized by SYBR Gold (Life Technologies, Waltham, MA) staining (Rodzen et al. 1998) in a scanner with multicolor fluorescence detection FMBIO III Plus (Hitachi, TYO, Japan). Allelic sizes were estimated by comparison with the sequencing ladder of SequiTherm™ Cycle Sequencing Kit (Epicentre Technologies, Madison, WI).

Data analyses

Allelic frequencies were obtained for each population and locus with GENETIX v.4.05 (Belkhir et al. 2004). Genetic diversity was evaluated by the number of alleles per locus (N_A) and by observed (H_O) and expected (H_E) heterozygosities. The effective number of alleles per locus (n_e) was estimated as the reciprocal of $\sum p_i^2$ (Hartl & Clark 1997). Genetic differences in the distribution of alleles among individuals from different locations, for each locus, were determined by Fisher's exact test (10,000 dememorization steps, 100 batches, and 10,000 iterations per batch) implemented in GENEPOP 4.2 (Raymond & Rousset 1995, Rousset 2008). Differences in H_O , H_E , and n_e among individuals from different locations were determined by Kruskal-Wallis with Stastitica 8. Hardy-Weinberg equilibrium (Markov chain, 1,000,000 steps, and 100,000 dememorization steps) and linkage disequilibrium (10,000 permutations) analyses were performed with Arlequin 3.1 (Excoffier et al. 2005). Using the Bonferroni approach, the significance level was then adjusted for multiple tests ($\alpha = 0.05$) (Rice 1989). The program HW-QuickCheck was used to assess the number of genotypes that present homozygote and heterozygote excess and heterozygote deficiencies at $P \leq 0.05$ (Kalinowski 2006).

Table 1. Genetic diversity of *Cassostrea corteziensis*, using three microsatellites at four locations: BL: Bahía de Lobos, LG: Las Glorias, BCe: Bahía de Ceuta, and BCa: Boca de Camichín. N : size sample; N_A : number of alleles; n_e : number of effective alleles; H_O : observed heterozygosities; H_E : expected heterozygosities; P : Fisher's exact test probabilities of Hardy-Weinberg equilibrium; f_{null} : estimate of null allele frequency. *Significant deviations from Hardy-Weinberg equilibrium ($P < 0.05$, ** $P < 0.001$) after sequential Bonferroni test.

Locus		BL	LG	BCe	BCa
<i>Ccor8</i>	N	50	42	48	40
	N_A	27	20	21	17
	n_e	13.889	8	11.211	7.459
	H_O	0.720	0.952	0.667	0.725
	H_E	0.937	0.886	0.920	0.877
	P	0.000**	0.025*	0.000**	0.002*
	f_{null}	0.106	0.000	0.122	0.058
<i>Ccor10</i>	N	45	43	50	42
	N_A	36	22	30	32
	n_e	16.946	7.625	10.730	13.466
	H_O	0.667	0.465	0.540	0.595
	H_E	0.952	0.879	0.916	0.937
	P	0.000**	0.000**	0.000**	0.000**
	f_{null}	0.141	0.211	0.188	0.166
<i>Ccor11</i>	N	47	39	50	43
	N_A	39	30	33	35
	n_e	27.962	19.133	24.038	25.156
	H_O	0.851	0.846	0.780	0.558
	H_E	0.975	0.960	0.968	0.972
	P	0.000**	0.000**	0.001*	0.000**
	f_{null}	0.059	0.050	0.094	0.205
<i>Ccor21</i>	N	47	38	36	41
	N_A	19	9	16	20
	n_e	3.181	2.216	4.516	4.625
	H_O	0.085	0.026	0.056	0.146
	H_E	0.693	0.556	0.790	0.793
	P	0.000**	0.000**	0.000**	0.000**
	f_{null}	0.362	0.350	0.409	0.357
<i>Ccor26</i>	N	43	33	40	36
	N_A	46	28	41	38
	n_e	34.240	19.276	28.319	24.452
	H_O	0.791	0.758	0.850	0.861
	H_E	0.982	0.963	0.977	0.973
	P	0.003*	0.000**	0.013*	0.011*
	f_{null}	0.091	0.096	0.054	0.050
<i>Ccor28</i>	N	48	43	47	43
	N_A	41	39	35	38
	n_e	27.926	26.604	28.692	23.553
	H_O	0.938	0.837	0.830	0.837
	H_E	0.974	0.974	0.976	0.969
	P	0.284	0.003*	0.000**	0.007*
	f_{null}	0.011	0.064	0.070	0.057
Mean	N	46.67	39.67	45.17	40.83
	N_A	34.67	24.67	29.5	30
	n_e	20.7	13.81	17.92	16.45
	H_O	0.675	0.647	0.621	0.620
	H_E	0.919	0.870	0.924	0.920

Estimation of larval retention and local replenishment was obtained by examining the genetic relatedness between pairs of individuals using Queller & Goodnight's estimator (1989); the statistical significance of this estimation was evaluated with 999 permutations and 1,000 bootstraps for each location, implemented in the software GENALEX (Peakall & Smouse 2012).

Frequencies of null alleles at each locus and population were estimated using the maximum expectation algorithm (Dempster et al. 1977) implemented in FreeNA (Chapuis & Estoup 2007).

Subsequent population genetic structure analyses were carried out as follows: 1) using original allele frequencies of all loci, 2) excluding allele frequencies of *Ccor21* locus, which showed a high frequency of null alleles, and 3) using allele frequencies after adjustment with FreeNA, when it was possible.

The fixation index F_{ST} (Weir & Cockerham 1984) was obtained to determine whether there were genetic differences among individuals from different locations, and global analysis of molecular variance (AMOVA) was performed. The P -value for the global AMOVA analysis used 50,175 permutations calculated in Arlequin 3.1 (Excoffier et al. 2005). A global F_{ST} analysis with a dataset carrying null alleles and using the excluding null alleles (ENA) correction method, which corrects the positive bias induced by the presence of null alleles in estimating F_{ST} , was also obtained. A 95% confidence interval for global F_{ST} was obtained with FreeNA, using 10,000 bootstraps (Chapuis & Estoup 2007). Genetic differences in allele frequencies between samples were also tested by paired F_{ST} analysis with original data, without *Ccor21* locus, and using the ENA correction method. P -values for paired F_{ST} analysis were obtained after 10,100 permutations calculated with Arlequin 3.1 (Excoffier et al. 2005), and 95% confidence intervals were obtained using 10,000 bootstraps with FreeNA (Chapuis & Estoup 2007). To assess the possible population structure among regional aggregates, a hierarchical AMOVA was performed on three groups: Bahía de Lobos (1), Las Glorias and Bahía de Ceuta (2), and Boca de Camichín (3). With hierarchical AMOVA differences at three levels were assessed: among regional groups relative to the total population (F_{CT}), among subpopulations relative to the regional groups (F_{SC}) and within subpopulations relative to the total population (F_{ST}).

Discriminant analysis of principal components (DAPC) was performed with the Adegnet package (Jombart 2008) in R (R Development Core Team 2011)

to explore genetic groups (Jombart & Collins 2021). First, the data were transformed through a PCA, and then a DAPC was applied (Jombart & Collins 2021). The best number of K clusters retaining 200 PCs was three, as indicated by the Bayesian information criterion (BIC). DAPC was run with 90 PCs and three DA eigenvalues, accounting for 90% of the variance.

Correlations between genetic and geographic distances were assessed to determine if oyster populations followed the isolation-by-distance model, using the Mantel test (100,000 permutations) with GENETIX (Belkhir et al. 2004). Reynolds genetic distances between locations were obtained with GENETIX, and geographic distances were measured as the shortest direct distances across the ocean between locations using Google Maps and a distance tool (<http://www.sunearthtools.com/es/tools/distance.php>).

Gene flow estimates were obtained using the private alleles: $Nm = e^{((\ln(pl)+2.44)/-0.505)}$ (Slatkin 1985), where pl was the mean frequency of private alleles. Also, an estimate of gene flow was obtained using equation 4 in Wang (2004), replacing F_{ST} obtained without *Ccor21* and the number of subpopulations.

First-generation migrants were evaluated to evaluate the potential recent gene flow using a Bayesian approach implemented in GENECLASS2 (Piry et al. 2004). Ten thousand individuals were simulated using a resampling algorithm (Rannala & Mountain 1997), with an alpha value of 0.01 and the statistical index L_{home}/L_{max} (Paetkau et al. 2004). L_{home} is the likelihood that an individual genotype belongs to the population from which it was sampled, given the observed allele frequencies. L_{max} is the maximum likelihood observed for this genotype in any population (Paetkau et al. 2004).

To assess the possible evidence of a recent bottleneck in the genetic disequilibrium of H_E , in addition to the presence of null alleles, Bottleneck software was included (<https://www1.montpellier.inrae.fr/CBGP/software/Bottleneck/bottleneck.html>). This software implemented the Two Phased Model (TPM), mode shift, Wilcoxon test, and 10,000 replications (Cornuet & Luikart 1997).

RESULTS

Genetic diversity

All loci showed differences among locations in allelic frequencies at the two most common alleles (except for the *Ccor21* locus). In the case of *Ccor21*, the most frequent allele was the same at all locations (Fig. S1).

Genetic diversity in *C. corteziensis* at six microsatellite loci ranged from moderate to high (9-46 for N_A , 1.027-21 for n_e , 0.026-0.952 for H_O , and 0.556-0.982 for H_E , Table 1). The mean genetic diversity estimator N_A indicated that Bahía de Lobos presented the highest diversity, whereas the lowest diversity was observed at Las Glorias (Table 1). The alleles distributions showed significant genic differentiation among all locations ($P < 0.008$), but no differences for H_O ($P = 0.9166$), H_E ($P = 0.6444$), or n_e ($P = 0.9166$) were found. All loci showed significant deviations from Hardy-Weinberg equilibrium (except in *Ccor28* at Bahía de Lobos), and all were caused by heterozygote deficiencies (except in *Ccor8* at Las Glorias). When analyzing the heterozygote deficiencies in HWQuickCheck, there was observed: 1) a lot of heterozygote deficiencies which rarely had a statistical significance, almost all were registered as ns (non-significant); 2) a low number of significant heterozygote excesses, and homozygotes excesses ($P < 0.05$) (not shown). Such differences between observed and expected genotypes contributed to the results obtained by the Hardy-Weinberg equilibrium test (Table 1). A tendency was observed in the mean heterozygote deficiencies among locations, with Bahía de Ceuta being the locality with the highest deficiency (Table 2).

At all locations, two loci showed significant linkage disequilibrium (*Ccor10-Ccor21*, $P < 0.0036$). However, we included both loci in the population structure analyses because linkage disequilibrium could be a combined effect of the low n_e at both loci (Table 1), similar frequencies at some alleles (Fig. S1), and the highest presence of null alleles in *Ccor21*.

Three locations (Bahía de Lobos, Bahía de Ceuta, and Boca de Camichín) showed mean relatedness coefficients (r) ranging from -0.02 to 0.02. Las Glorias

Table 2. Number of genotypes with heterozygote deficiencies, using six microsatellites at four locations: BL: Bahía de Lobos, LG: Las Glorias, BCe: Bahía de Ceuta, and BCa: Boca de Camichín.

Locus	BL	LG	BCe	Bca
<i>Ccor8</i>	170	76	108	89
<i>Ccor10</i>	185	88	155	153
<i>Ccor11</i>	306	212	303	252
<i>Ccor21</i>	48	35	103	74
<i>Ccor26</i>	198	175	262	208
<i>Ccor28</i>	276	201	287	228
Mean	197	131	203	167

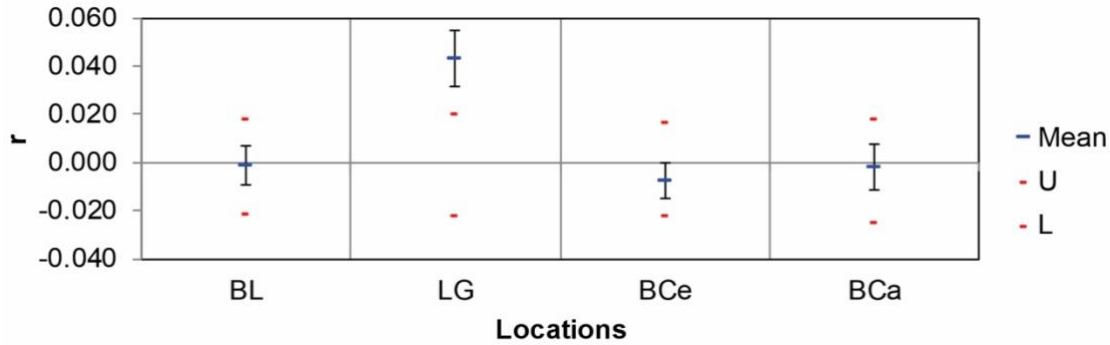


Figure 2. Mean relatedness coefficient (r) between individuals in the same locality in blue lines. Red lines represent the upper (U) and lower (L) confidence limits bound by the 95% confidence interval about the null hypothesis of no difference across the populations as determined by permutation. Black lines show upper and lower error bars bound by the 95% confidence interval about the mean values determined by bootstrap resampling. Locations on the east coast of the Gulf of California include LB: Bahía Lobos, LG: Las Glorias, BCe: Bahía Ceuta, and BCa: Boca de Camichín.

showed a significantly high value ($r = 0.043$, $P = 0.004$), indicating that individuals were more similar from the same location than expected under random mating (Fig. 2).

Null alleles were present for all loci from all locations (except in *Ccor8* at Las Glorias), with frequencies from 5-21.1%, except for *Ccor21*, with frequencies from 35-40.9% (Table 1). Subsequent analyses, including and excluding allele frequencies of the *Ccor21* locus, were done due to the possible bias of this locus in the resulting genetic population structure.

Population genetic structure

The global AMOVA showed low genetic differentiation among populations when all loci were included ($F_{ST} = 0.0135$), and after *Ccor21* was excluded ($F_{ST} = 0.0111$), both F_{ST} values were highly significant from zero ($P = 0.0000$). The pairwise F_{ST} analysis with all loci showed significant differences between all locations, except between Bahía de Lobos and Bahía de Ceuta ($P = 0.1096$) and between Bahía de Ceuta and Boca de Camichín ($P = 0.0494$). Las Glorias showed genetic differentiation with the other three locations. After excluding the *Ccor21* locus from pairwise F_{ST} analysis, all comparisons between locations were significant, except between Las Glorias and Bahía de Ceuta (Table 3). Similar values of global F_{ST} were obtained, including all loci and using *ENA* correction ($F_{ST} = 0.0097$, CI: 0.0026-0.0219) or not ($F_{ST} = 0.0102$, CI: 0.0031-0.0223), as well as pairwise F_{ST} (Fig. S2). A hierarchical AMOVA with all loci did not show significant differences among regional groups relative to the total population ($F_{CT} = 0.0027$, $P = 0.3309$). Still, there were significant differences among subpopu-

lations relative to the regional groups ($F_{SC} = 0.0114$, $P = 0.0123$) and within subpopulations relative to the total population ($F_{ST} = 0.0140$, $P = 0.0000$). Hierarchical analysis without *Ccor21* locus did not show significant differences between regional groups relative to the total population ($F_{CT} = 0.0073$, $P = 0.1663$) or between subpopulations relative to the regional groups ($F_{SC} = 0.0051$, $P = 0.0669$). Still, highly significant differentiation was observed within subpopulations relative to the total population ($F_{ST} = 0.0123$, $P = 0.0000$).

The DAPC and pairwise F_{ST} without the locus *Ccor21* showed similar results (Table 3). Three genetic groups were observed: Bahía de Lobos, Boca de Camichín, and a third group composed of Las Glorias and Bahía de Ceuta (Fig. 3). These two locations are the geographically nearest of the four. However, the subpopulations of pleasure oysters found in this investigation did not agree with the isolation-by-distance model because no significant correlation among genetic and geographic distances was found with all loci ($r = 0.037$, $P = 0.333$) or without *Ccor21* locus ($r = 0.137$, $P = 0.542$).

Estimates of gene flow obtained using private alleles showed a gradient from north to south, where Bahía de Lobos showed the lowest estimate ($Nm = 17$), and Las Glorias ($Nm = 25$), Bahía de Ceuta ($Nm = 27$), and Boca de Camichín ($Nm = 28$) showed similar estimates. The Nm obtained with the equation 4 in Wang was lower than the previous ($Nm = 12.5$).

The results of GENECLASS2 identified 12 first-generation migrants ($P < 0.01$) out of 188 individuals (6.4%). Asymmetric gene flow occurred, where the main source of migrants was Boca de Camichín (BCa)

Table 3. Pairwise F_{ST} between samples from BL: Bahía de Lobos, LG: Las Glorias, BCe: Bahía de Ceuta, and BCa: Boca de Camichín, including all loci (above diagonal) and without *Ccor21* (below diagonal). *Significant F_{ST} P -values ($P < 0.025$); ** $P < 0.001$, after sequential Bonferroni test.

	BL	LG	BCe	BCa
BL	--	0.0128*	0.0053	0.0113*
LG	0.0139**	--	0.0101*	0.0185**
BCe	0.0060 *	0.0056	--	0.0078
BCa	0.0114**	0.0189**	0.0128**	--

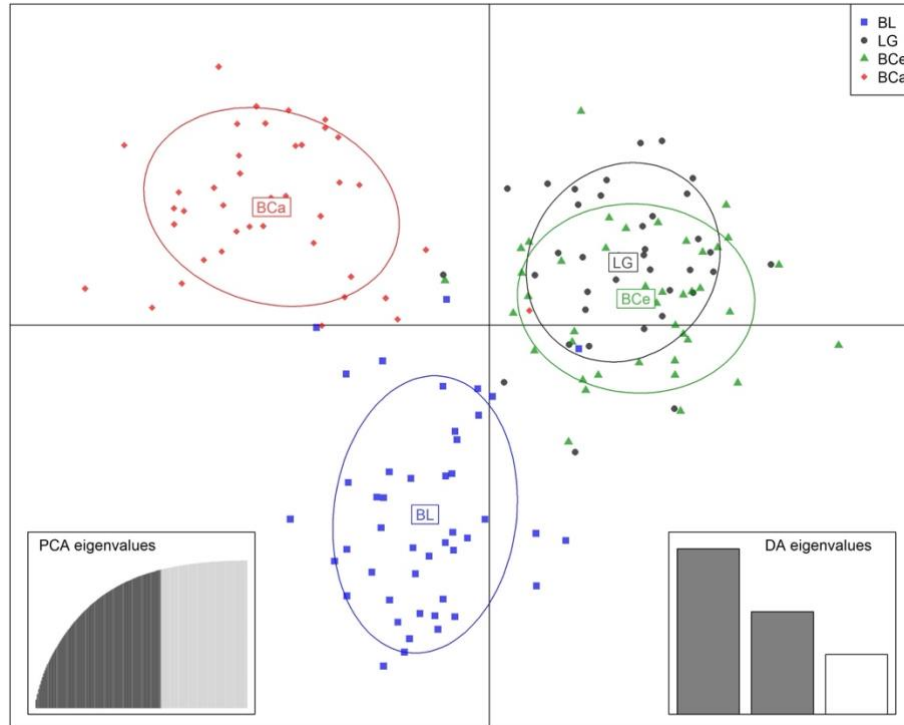


Figure 3. Discriminant analysis of principal components (DAPC) of *Crassostrea corteziensis* from four sampled sites based on five microsatellite loci (without *Ccor21*), 90 PCAs, and three DA eigenvalues. Sampled locations include BL: Bahía de Lobos, LG: Las Glorias, BCe: Bahía de Ceuta, and BCa: Boca de Camichín.

with five migrants (2.6%), followed by Bahía de Lobos (BL) with four migrants (2%). In addition, the most important sink site was Bahía de Ceuta, receiving eight migrants (4.2%) from BL and BCa (Fig. 4). Results with Bottleneck software showed no evidence of a recent bottleneck at any location.

DISCUSSION

Significant differences in the distribution of alleles, global F_{ST} , pairwise F_{ST} , and DAPC strongly suggest the existence of a non-panmictic population of *C. corteziensis* along the study area. The levels of genetic

differentiation among subpopulations of *C. corteziensis* ($F_{ST} = 0.0097$ – 0.0135) were similar to those reported for *Ostrea edulis* and *C. virginica* ($F_{ST} = 0.019$, Launey et al. 2002, Galindo-Sánchez et al. 2008), but lower to that reported for *C. virginica* with SNPs ($F_{ST} = 0.043$, Varney et al. 2009). Our observations differ from those of Pérez-Enríquez et al. (2008), who reported no genetic structure and high gene flow in *C. corteziensis* for the same study area. Uniformity in allozyme frequencies, suggesting balancing selection, can decrease population differentiation by maintaining the same alleles in different populations, operating concerning the observed electromorph classes instead of the level

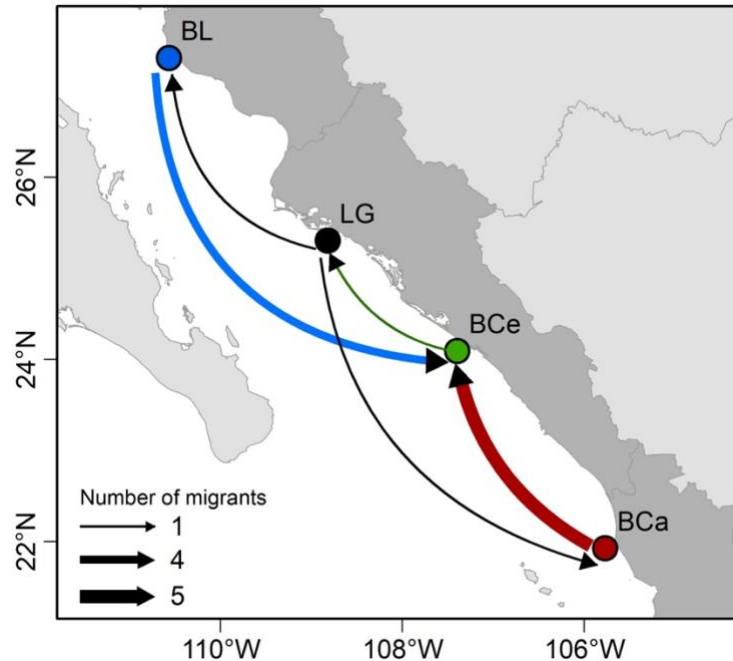


Figure 4. Spatial network from first-generation migrants. The width of the lines is scaled according to the number of migrants, and the color and direction of the arrows represent the location from which they came out. BL: Bahía de Lobos, LG: Las Glorias, BCe: Bahía de Ceuta, and BCa: Boca de Camichín.

of the hidden variation (Karl & Avise 1992, Freeland 2007).

In the present study, all loci were in disequilibrium of Hardy-Weinberg caused by a deficit of heterozygotes. Several causes for deviations from Hardy-Weinberg equilibrium have been inferred: null alleles, the Wahlund effect, selection against heterozygotes, non-random mating, genotyping errors, and inbreeding (McGoldrick et al. 2000, Rose et al. 2006, Galindo-Sánchez et al. 2008, Enríquez-Espinoza & Grijalva-Chon 2010, Chen et al. 2017, Jiang et al. 2023). The best explanation for deviations from Hardy-Weinberg equilibrium in the present study was the presence of null alleles in all loci.

After *ENA* corrections with all loci, the global F_{ST} of *C. corteziensis* was almost the same. Pairwise F_{ST} values remained very similar, as already reported by Galindo-Sánchez et al. (2008) and Varney et al. (2009), who, after reanalyzing the data to account for null alleles, found that the global F_{ST} value only differed slightly from the initial value. However, after excluding *Ccor21*, pairwise F_{ST} values showed significant differences in all comparisons except Las Glorias and Bahía Ceuta. This result underlined the strong bias of *Ccor21*, distorting the perceived levels of population

differentiation. For this reason, the discussion will be carried out without *Ccor21*.

Genetic population differentiation between populations of *C. corteziensis* on the east coast of the Gulf of California could be the result of several factors, like random genetic drift, variance in reproductive success, the short pelagic phase, larval retentions in protected areas (estuaries along the coast), local replenishment (local gene flow), asymmetric and restricted gene flow, and local adaptations to temperature, salinity, and other environmental conditions within each location.

Random genetic drift in finite local populations maintained by larval retention was reported in *C. gigas* and *C. virginica*. Partial isolation of major estuarine populations would help explain such larval retention (Hedgecock 1994).

Variance in reproductive success may be large enough to limit effective population numbers to fractions of actual abundances. Substantial variation in reproductive success is made possible by great fecundity (60×10^6 eggs in *C. corteziensis*) and high early mortality (type III survivorship curves). A small minority of individuals can replace each generation's entire population by sweepstakes-chance matching reproductive activity with oceanographic conditions

conducive to spawning, fertilization, larval survival, and successful recruitment (Hedgecock 1994).

The pleasure oyster shows at least two strong spawning periods, one during the summer and the second in autumn (Stuardo & Martínez 1975, Rodríguez-Jaramillo et al. 2008). Under summer circulation conditions, from June to September (Marinone 2012, Santiago-García et al. 2014), larvae of *C. corteziensis* could be dispersed on the continental shelf over hundreds of kilometers towards the north. During autumn circulation conditions (from September to December), retention and dispersal of *C. corteziensis* larvae towards the south could occur (Santiago-García et al. 2014). In October, retention could contribute to the population structure found in this study. During November and December, during a shorter period of oceanographic conditions than in summer, larvae dispersal could occur towards the south.

However, genetic connectivity not only involves dispersal from one place to another but also the survival and reproduction of migrants in such a way that they contribute to the local genetic pool perpetuated for subsequent generations (Hedgecock et al. 2007). During its short larval period of 18-22 days, oyster larvae coming from an estuary after dispersal could fail to find an adequate habitat to settle, like mangrove roots or fixed substrate in distant estuaries. Genetic differentiation on a local scale requires differential survival of genotypes after recruitment or temporal variation in the genetic composition of recruits. Although Johnson & Black (1984) believe that temporal variance arises from temporally and spatially varying selection on larvae, an alternative explanation is sampling variance due to the sweepstakes reproductive success of a small fraction of the population (Hedgecock 1994).

In the case of Las Glorias, this location showed the lowest number of alleles, the lowest mean expected heterozygosity and a significant mean of relatedness. Such results and the hierarchical AMOVA suggest that this sample was not randomly collected from a wild population. Because this sample was bought by an oyster restaurateur who had grown these oysters in an intertidal enclosure and stated that the sample was wild, the possibility that these individuals came from a reduced number of progenitors cannot be ruled out. In *Saccostrea echinata*, relatively low haplotype and nucleotide diversity indices were detected in one location. Individuals from this sample were collected from an oyster farm and may have been from a single or few cohort(s) (Nowland et al. 2019). Oysters are highly fecund, and hatcheries sometimes use a few

broodstock to produce offspring, with consequences for heterozygosity.

According to the genetic structure results (F_{ST} and DAPC), three populations were supported: 1) Boca de Camichín at the entrance of the Gulf of California, 2) Bahía de Ceuta and Las Glorias in the southern Gulf of California, and 3) Bahía de Lobos in the southern Gulf of California. Bahía de Ceuta and Las Glorias, the two nearest locations in the middle, were the most similar. Besides, first-generation migrants indicated asymmetric gene flow between locations, showing Bahía de Ceuta as a sink population, Bahía de Lobos and Boca de Camichín as source populations, and Las Glorias as a source-sink population to a lesser extent. Such genetic structure exists in the presence of gene flow because of the asymmetry in the gene flow. Given the two reproductive cycles in *C. corteziensis* (summer and autumn), the larval dispersal was driven by the currents of the seasonal circulation in the Gulf of California. Variance in reproductive success can also affect population structure because of temporal and spatial variation in the genetic composition of recruiting larvae (Hedgecock 1994). This observation was supported by previous studies in *Spondylus calcifer*, *Panopea globosa*, and *Callinectes bellicosus* (Soria et al. 2012, Munguia-Vega et al. 2015, Cisneros-Mata et al. 2018). In addition, all locations showed heterozygotes deficiencies, explained to a good extent by the presence of null alleles, but the existence of genotypes (heterozygotes and homozygotes) in higher frequencies than expected in Hardy-Weinberg equilibrium suggests the hypothesis of genotypes selection under different environmental conditions.

Understanding a genetic stock is important for managing fisheries better before aquaculture occurs. Natural populations contain standing genetic variation, which is critical for adaptation to changing environments (Barrett & Schluter 2008). Oyster populations identified as genetic groups must be considered evolutionary significant units and should be managed with conservation strategies (Lapègue et al. 2023).

The results of this work increased our understanding of the distribution of the pleasure oyster genetic groups. Assessing genetic diversity is beneficial for better protection and use of local germplasm for the development of oyster farming in the same location. In the future, a new study with SNPs that could confirm or discard the results obtained here with microsatellites is needed. In the case of SNPs, local adaptations to environmental conditions could be explored with outlier loci.

Credit author contribution

N. Díaz-Viloria: conceptualization, validation, methodology, formal analysis, writing-original draft, review, and editing; A.M. Ibarra: Funding acquisition, project administration, supervision, review, and editing; C.E. Rivera-Apodaca: methodology and formal analysis; N. Reguera-Rouzaud: data curation, formal analysis, review, and editing; N.K. Hernández-Ibarra: formal analysis, review, and editing; P. Cruz-Hernández: conceptualization, methodology, formal analysis, validation, supervision, review, and editing. All authors have read and accepted the published version of the manuscript.

Conflict of interest

The authors declare no conflict of interest.

ACKNOWLEDGMENTS

We are indebted to Susana Ávila Álvarez and José Luis Ramírez Arce of CIBNOR for technical assistance. This work was funded by grant No. 2003-02-035 from SAGARPA-CONACYT to A.M. Ibarra. Claudia Gabriela Rivera-Apodaca was supported by a graduate fellowship from CONACYT during her M.Sc. studies at CICIMAR.

REFERENCES

- Barrett, R.D. & Schluter, D. 2008. Adaptation from standing genetic variation. *Trends in Ecology & Evolution*, 23: 38-44. doi: 10.1016/j.tree.2007.09.008
- Belkhir, K., Borsa, P., Chikhi, L., Raufaste, N., et al. 2004. Genetix 4.05 logiciel sous Windows TM pour la génétique des populations. Université de Montpellier II, Montpellier. [http://www.univ-montp2.fr/~genetix/genetix/genetix.htm]. Reviewed: December 5, 2023.
- Chapuis, M.P. & Estoup, A. 2007. Microsatellite null alleles and estimation of population differentiation. *Molecular Biology and Evolution*, 24: 621-631. doi: 10.1093/molbev/msl191
- Chávez-Villalba, J., López-Tapia, M., Mazon-Suástegui, J.M. & Robles-Mungaray, M. 2005. Growth of the oyster *Crassostrea corteziensis* (Hertlein, 1951) in Sonora, México. *Aquaculture Research*, 36: 1337-1344. doi: 10.1111/j.1365-2109.2005.01345.x
- Chen, B., Cole, J.W. & Grond-Ginsbach, C. 2017. Departure from Hardy Weinberg Equilibrium and genotyping error. *Frontiers in Genetics*, 8: 167. doi: 10.3389/fgene.2017.00167
- Cisneros-Mata, M.A., Munguía-Vega, A., Rodríguez-Félix, D., Aragón-Noriega, E.A., et al. 2018. Genetic diversity and metapopulation structure of the brown swimming crab (*Callinectes bellicosus*) along the coast of Sonora, México: implications for fisheries management. *Fisheries Research*, 212: 97-106. doi: 10.1016/j.fishres.2018.11.021
- Cornuet, J.M. & Luikart, G. 1997. Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics*, 144: 2001-2014. doi: 10.1093/genetics/144.4.2001
- Cruz, P., Yáñez-Jacome, B., Ibarra, A.M. & Rangel-Becerril, J. 2007. Isolation and characterization of microsatellite loci in the Pacific pleasure oyster, *Crassostrea corteziensis*, and their cross-species amplification in four other oyster species. *Molecular Ecology Notes*, 7: 448-450. doi: 10.1111/j.1471-8286.2006.01613.x
- Dempster, A.P., Laird, N.M. & Rubin, D.B. 1977. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society*, 39: 1-38.
- Enriquez-Espinoza, T.L. & Grijalva-Chon, J.M. 2010. Genetic variability of *Crassostrea gigas* and *Crassostrea corteziensis* from a hatchery in northwestern México. *Ciencias Marinas*, 33: 333-344. doi: 10.7773/cm.v36i4.1609
- Excoffier, L., Laval, G. & Schneider, S. 2005. Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, 1: 47-50.
- Freeland, J.R. 2007. *Molecular ecology*. John Wiley & Sons, Hoboken.
- Galindo-Sánchez, C.E., Gaffney, P.M., Pérez-Rostro, C.I., De la Rosa-Vélez, J., et al. 2008. Assessment of genetic diversity of the eastern oyster *Crassostrea virginica* in Veracruz, México using microsatellite markers. *Journal of Shellfish Research*, 27: 721-727. doi: 10.2983/0730-8000(2008)27[721:AOGDOT]2.0.CO;2
- Hartl, D.L. & Clark, A.G. 1997. *Principles of population genetics*. Sinauer Associates, Sunderland.
- Hedgecock, D. 1994. Does variance in reproductive success limit effective population sizes of marine organisms? In: Beaumont, A.R. (Ed.). *Genetics and evolution of aquatic organisms*. Chapman & Hall, London, pp. 122-135.
- Hedgecock, D., Barger, P.H. & Edmands, S. 2007. Genetic approaches to measuring connectivity. *Oceanography*, 20: 70-79. doi: 10.5670/oceanog.2007.30
- Hellberg, M.E., Burton, R.S., Neigel, J.E. & Palumbi, S.R. 2002. Genetic assessment of connectivity among marine populations. *Bulletin of Marine Science*, 70: 273-290.
- Jiang, G., Zhang, Y., Du, L., Chen, Y., et al. 2023. Genetic diversity and structure in a selected strain of hybrid

- oysters between *Crassostrea gigas* and *C. angulata* evaluated from microsatellites and mitochondrial COI sequences. *Aquaculture*, 574: 739716. doi: 10.1016/j.aquaculture.2023.739716
- Johnson, M.S. & Black, R. 1984. Pattern beneath the chaos: the effect of recruitment on genetic patchiness in an intertidal limpet. *Evolution*, 38: 1371-1383. doi: 10.1111/j.1558-5646.1984.tb05658.x
- Jombart, T. 2008. Adegnet: A R package for the multivariate analysis of genetic markers. *Bioinformatics*, 11: 1403-1405. doi: 10.1093/bioinformatics/btn129
- Jombart, T. & Collins, C. 2021. A tutorial for discriminant analysis of principal components (DAPC) using Adegnet 2.1.3. [<https://adegenet.r-forge.rproject.org/files/tutorial-dapc.pdf>]. Reviewed: April 19, 2024.
- Kalinowski, S.T. 2006. Quickcheck: an easy-to-use computer program for checking genotypes to ensure agreement with Hardy-Weinberg expectations. *Molecular Ecology Notes*, 6: 974-979. doi: 10.1111/j.1471-8286.2006.01456.x
- Karl, A. & Avise, J.C. 1992. Balancing selection at allozyme loci in oysters: Implications from nuclear RFLPs. *Science*, 256: 100-102. doi: 10.1126/science.1348870
- Lapègue, S., Reisser, C., Harrang, E., Heurtebise, S., et al. 2023. Genetic parallelism between European flat oyster populations at the edge of their natural range. *Evolutionary Applications*, 16: 393-407. doi: 10.1111/eva.13449
- Launey, S., Ledu, C., Boudry, P., Bonhomme, F., et al. 2002. Geographic structure in the European flat oyster (*Ostrea edulis* L.) as revealed by microsatellite polymorphism. *Journal of Heredity*, 93: 331-338. doi: 10.1093/jhered/93.5.331
- Lavín, M.F. & Marinone, S.G. 2003. An overview of the physical oceanography of the Gulf of California. In: Velasco-Fuentes, O.U., Sheinbaum, J. & Ochoa, J. (Eds.). *Nonlinear processes in geophysical fluid dynamics*. Springer, Berlin, pp. 173-204.
- Marinone, S.G. 2003. A three-dimensional model of the mean and seasonal circulation of the Gulf of California. *Journal of Geophysical Research*, 108: 3325. doi: 10.1029/2002JC001720
- Marinone, S.G. 2012. Seasonal surface connectivity in the Gulf of California. *Estuarine, Coastal and Shelf Science*, 100: 133-141. doi: 10.1016/j.ecss.2012.01.003
- McGoldrick, D.J., Hedgecock, D., English, L.J., Baoprasertkul, P., et al. 2000. The transmission of microsatellite alleles in Australian and North American stocks of the Pacific oyster (*Crassostrea gigas*): selection and null alleles. *Journal of Shellfish Research*, 19: 779-788.
- Munguia-Vega, A., Leyva-Valencia, I., Lluch-Cota, D.B. & Cruz-Hernandez, P. 2015. Genetic structure of the Cortes geoduck *Panopea globosa* Dall, 1898, from the Mexican northwest. *Journal of Shellfish Research*, 34: 153-161. doi: 10.2983/035.034.0119
- Nowland, S.J., Silva, C.N.S., Southgate, P.C. & Strugnelli, J.M. 2019. Mitochondrial and nuclear genetic analyses of the tropical black-lip rock oyster (*Saccostrea echinata*) reveal population subdivision and inform sustainable aquaculture development. *BMC Genomics*, 20: 711. doi: 10.1186/s12864-019-6052-z
- Paetkau, D., Slade, R., Burden, M. & Estoup, A. 2004. Genetic assignment methods for the direct, real-time estimation of migration rate: a simulation-based exploration of accuracy and power. *Molecular Ecology*, 13: 55-65. doi: 10.1046/j.1365-294x.2004.02008.x
- Peakall, R. & Smouse, P.E. 2012. GenALEX 6.5: Genetic analysis in Excel. Population genetic software for teaching and research -an update. *Bioinformatics*, 28: 2537-2539. doi: 10.1093/bioinformatics/bts460
- Pérez-Enríquez, R., Ávila, S. & Ibarra, A.M. 2008. Population genetics of the oyster *Crassostrea corteziensis* in the Gulf of California. *Ciencias Marinas*, 34: 479-490.
- Piry, S., Alapetite, A., Cornuet, J.M., Paetkau, D., et al. 2004. GENECLASS2: a software for genetic assignment and first-generation migrant detection. *Journal of Heredity*, 95: 536-539. doi: 10.1093/jhered/esh074
- Queller, D.C. & Goodnight, K.F. 1989. Estimating relatedness using genetic markers. *Evolution*, 43: 258-275. doi: 10.1111/j.1558-5646.1989.tb04226.x
- R Development Core Team. 2011. R A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. [<https://www.r-project.org/>]. Reviewed: April 19, 2024.
- Rannala, B. & Mountain, J.L. 1997. Detecting immigration by using multilocus genotypes. *Proceedings of the National Academy of Sciences*, 94: 9197-9201. doi: 10.1073/pnas.94.17.9197
- Raymond, M. & Rousset, F. 1995. Genepop (versión 3.4): Populations genetics software for exact tests and ecumenicism. *Journal of Heredity*, 86: 248-249.
- Rice, W.R. 1989. Analyzing tables of statistical tests. *Evolution*, 43: 223-225. doi: 10.2307/2409177
- Rodríguez-Jaramillo, C., Hurtado, M.A., Romero-Vivas, E., Ramírez, J.L., et al. 2008. Gonadal development and histochemistry of the tropical oyster, *Crassostrea corteziensis* (Hertlein, 1951) during an annual reproductive cycle. *Journal of Shellfish Research*, 27: 1129-1141. doi: 10.2983/0730-8000-27.5.1129
- Rodríguez-Romero, F., García-Saez, C. & Laguarda-Figueras, A. 1988. Electrophoretic patterns variation

- in two oyster populations of *Crassostrea corteziensis* from the Mexican coast. *Anales del Instituto de Ciencias del Mar y Limnología*, 15: 177-184.
- Rodzen, J.A., Agresti, J.J., Tranah, G. & May, B. 1998. Agarose overlays allow simplified staining of polyacrylamide gels. *Biotechniques*, 25: 584. doi: 10.2144/98254bm07
- Rose, C.G., Kennedy, T.P. & Matthew, P.H. 2006. Isolation by distance in the Eastern oyster, *Crassostrea virginica*, in Chesapeake Bay. *Journal of Heredity*, 97: 158-170. doi: 10.1093/jhered/esj019
- Rousset, F. 2008. Genepop'007: A complete reimplementation of the Genepop software for Windows and Linux. *Molecular Ecology Resources*, 8: 103-106. doi: 10.1111/j.1471-8286.2007.01931.x
- Santiago-García, M.W., Marinone, S.G. & Velasco-Fuentes, O.U. 2014. Three-dimensional connectivity in the Gulf of California based on a numerical model. *Progress in Oceanography*, 123: 64-73. doi: 10.1016/j.pocean.2014.02.002
- Silliman, K. 2018. Population structure, genetic connectivity, and adaptation in the Olympia oyster (*Ostrea lurida*) along the west coast of North America. *Evolutionary Applications*, 12: 923-939. doi: 10.1111/eva.12766
- Slatkin, M. 1985. Rare alleles as indicators of gene flow. *Evolution*, 1: 53-65. doi: 10.1111/j.1558-5646.1985.tb04079.x
- Soria, G., Munguía-Vega, A., Marinone, S.G., Moreno-Báez, M., et al. 2012. Linking bio-oceanography and population genetics to assess larval connectivity. *Marine Ecology Progress Series*, 463: 159-175. doi: 10.3354/meps09866
- Stuardo, J. & Martínez, A. 1975. Relaciones entre algunos factores ecológicos y la biología, de poblaciones de *Crassostrea corteziensis* Hertlein, 1951, de San Blas, Nayarit, México. *Anales del Instituto de Ciencias del Mar y Limnología*, 2: 89-130.
- Varney, R.L., Galindo-Sánchez, E., Cruz, P. & Gaffney, P.M. 2009. Population genetics of the eastern oyster *Crassostrea virginica* (Gmelin, 1791) in the Gulf of México. *Journal of Shellfish Research*, 28: 855-864. doi: 10.2983/035.028.0415
- Wang, J. 2004. Application of the one-migrant-per-generation rule to conservation and management. *Conservation Biology*, 18: 332-343. doi: 10.1111/j.1523-1739.2004.00440.x
- Weir, B.S. & Cockerham, C.C. 1984. Estimating F-statistics for the analysis of population structure. *Evolution*, 38: 1358-1370. doi: 10.2307/2408641

Received: June 24, 2024; Accepted: October 25, 2024

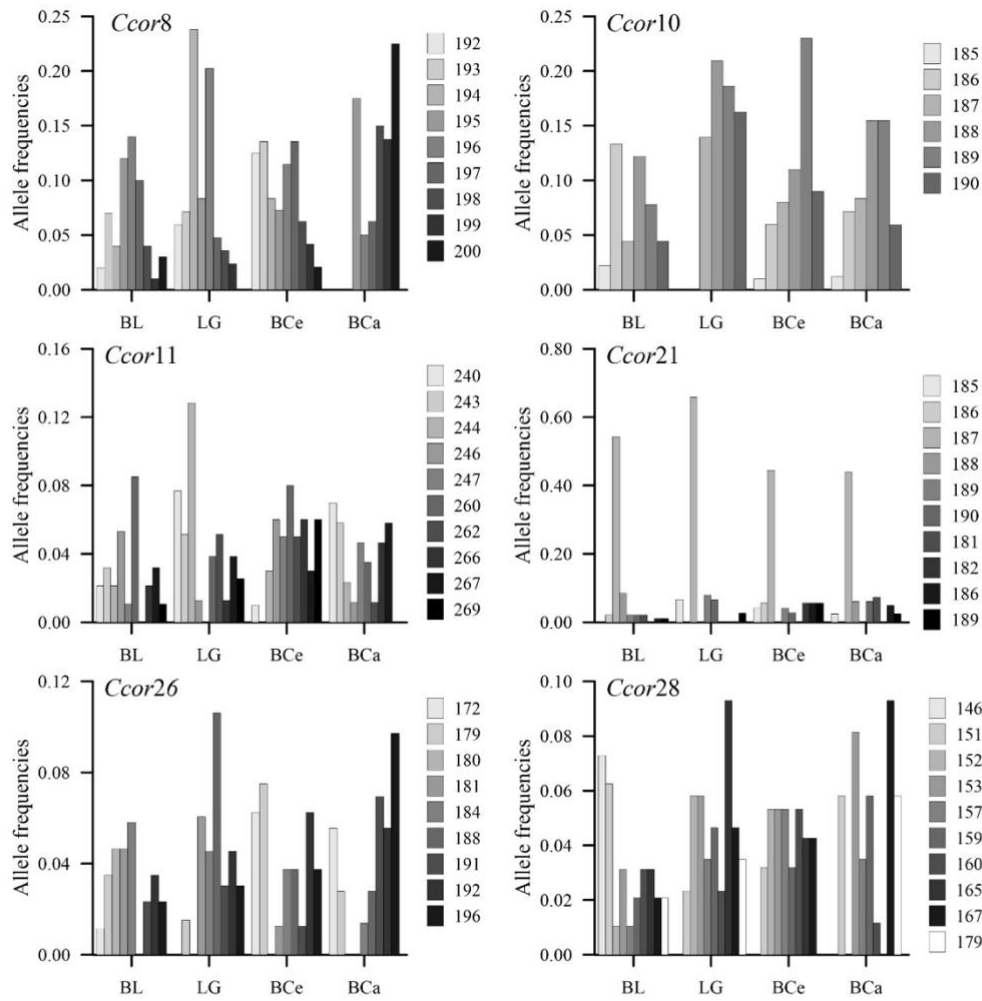


Figure S1. Allele frequencies of *Crassostrea corteziensis*. BL: Bahía de Lobos, LG: Las Glorias, BCe: Bahía de Ceuta, and BCa: Boca de Camichín. Only major alleles are shown (on the right side of each graph).

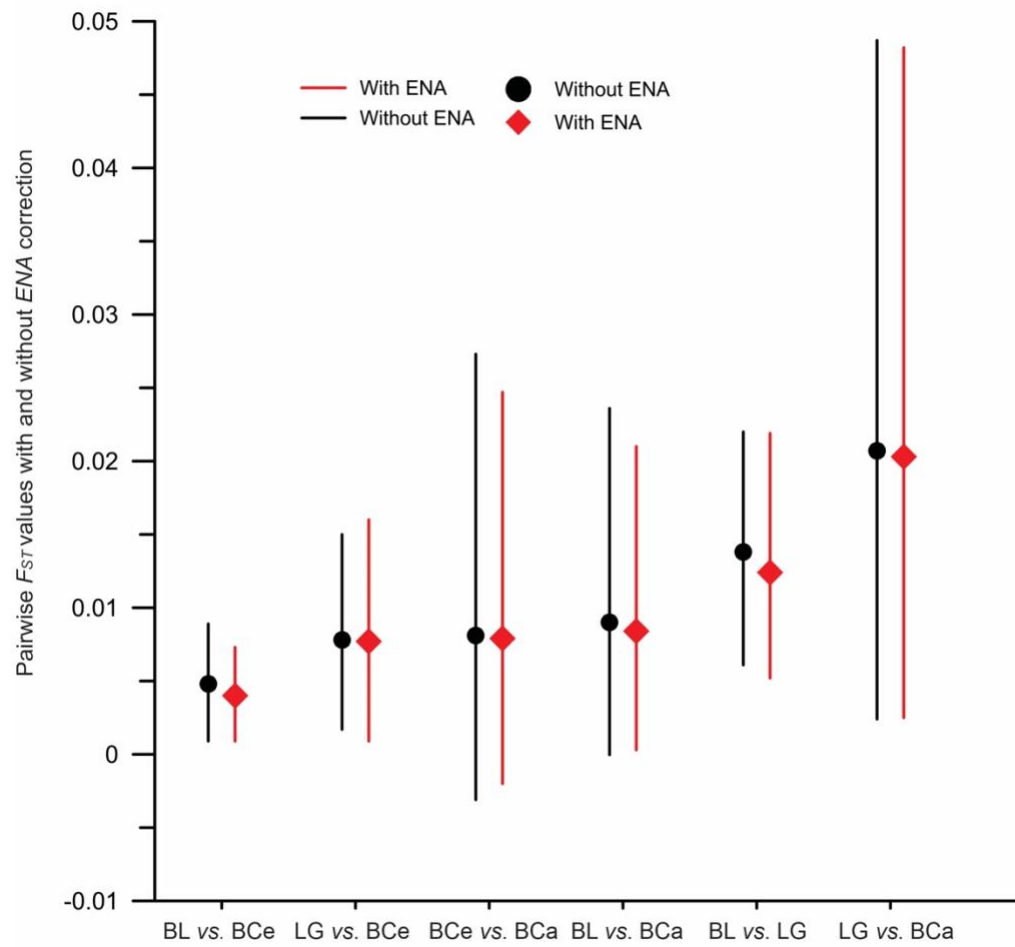


Figure S2. Pairwise F_{ST} (10,100 permutations) with and without *ENA* correction and 95% confidence intervals (10,000 bootstraps) between samples from BL: Bahía de Lobos, LG: Las Glorias, BCe: Bahía de Ceuta, and BCa: Boca de Camichín.